Zeitreihenökonometrie

Folien

Rolf Tschernig & Christoph Knoppik

Universität Regensburg

Februar 2026[1]

[1]Hinweis: Diese Folien wurden in vorherigen Versionen mit *Folien für Ökonometrie II* betitelt. Zu den Änderungen der einzelnen Versionen siehe Appendix C. Wir danken herzlich Kathrin Kagerer, Joachim Schnurbus, Roland Weigand und Stefan Rameseder für ihre Korrekturen und Verbesserungsvorschläge. Wir danken Patrick Kratzer für die Übersetzung der EViews-Workfiles und Programme in R-Programme.

Kursbeschreibung

In der Veranstaltung **Zeitreihenökonometrie** steht die Analyse von Zeitreihendaten im Mittelpunkt. Die Analyse von Zeitreihendaten erfordert i.A. eine Abschwächung einiger in der Veranstaltung **Einführung in die Ökonometrie** eingeführten Voraussetzungen für das lineare Regressionsmodell, wodurch es nunmehr möglich ist, die Schätzeigenschaften für eine gegebene Stichprobengröße durch so genannte asymptotische Schätzeigenschaften zu approximieren. Dieses Vorgehen ist auch notwendig, wenn statistische Tests bei unbekannter Fehlervertei-

lung durchgeführt werden. Ferner wird aufgezeigt, auf welche Weise die asymptotischen Eigenschaften von den dynamischen Stabilitätseigenschaften des stochastischen Mechanismus abhängen, der die beobachteten Daten generiert haben könnte, und wie diese sog. Stationaritätseigenschaften für autoregressive Zeitreihenmodelle überprüft werden können. Eng verbunden hiermit ist die Fragestellung, ob ein beobachteter Zeittrend deterministischer Natur ist oder eine Ausprägung eines Random Walks ist. Zur Beantwortung dieser Frage sind Einheitswurzeltests notwendig. Im zweiten Teil der Veranstaltung werden multivariate Zeitreihenmodelle eingeführt, die die Analyse mehrerer Zeitreihen und vorliegender Interdependenzen erlauben. Prägen Zeittrends mehrere Zeitreihen gemeinsam, so liegen langfristige Gleichgewichtsbeziehungen vor, zu deren Modellierung Kointegrationsmodelle verwendet werden. Diese Modelle spielen insbesondere in der empirischen Makroökonomie eine herausragende Rolle und sind ein unverzichtbares

Werkzeug für Prognosen.

Inhaltsverzeichnis

A. R-Programme für die empirischen Beispiele I

Abbildungsverzeichnis

Tabellenverzeichnis

Organisation

Kontakt

Prof. Dr. Rolf Tschernig
Gebäude RW(L), 5. Stock, Raum 514
Universitätsstr. 31, 93040 Regensburg
Tel. (+49) 941/943 2737, Fax (+49) 941/943 4917
Email: rolf.tschernig@wiwi.uni-regensburg.de

https://www.uni-regensburg.de/wirtschaftswissenschaften/fakultaet/org/oekonometrie

Zeiten, Räume und Kursleiter

siehe Kurshomepage und dort angegebene Links zum Vorlesungsverzeichnis (SPUR)

https://www.uni-regensburg.de/wirtschaftswissenschaften/fakultaet/org/oekonometrie/lehre/bachelor/zeitreihenoekonometrie

Voraussetzungen

Inhalt des Moduls **Einführung in die Ökonometrie**.

Prüfungsmodalitäten und Notengebung

Zeitreihenökonometrie ist ein Wahlpflichtkurs in der B.Sc.-VWL-Schwerpunktmodulgruppe **Data Science for Economics** (PO 2024, 2021), der B.Sc.-BWL-Vertiefungsmodulgruppe **Themen der Volkswirtschaftslehre**, in der B.Sc.-Digital-Business-Pflichtmodulgruppe **Data Analytics** sowie in BA-Schwerpunktmodulgruppen. Weitere Informationen enthält der Modulkatalog

https://wiwi-portal.app.uni-regensburg.de/studyprogram/public.php

Auch kann das Modul als Alternative zu dem Modul **Zeitreihen** (105 DAT-B-ELM-TIME) im B.Sc. Data Science belegt werden, siehe Modulkatalog https://www.uni-regensburg.de/fileadmin/user_upload/bilderkatalog/ordnungen-satzungen-richtlinien/studium/modulbeschreibungen/bsc/BSc_Data_Science_MK_WiSe2526.pdf

Aktuelle Regelungen finden Sie auf GRIPS:

https://elearning.uni-regensburg.de/course/view.php?id=42624

Software

Im Kurs wird die freie Software **R** (http://www.r-project.org/) verwendet. In Appendix A sind alle in R-Programme aufgelistet, die in diesen Folien für empirische Analysen oder Monte-Carlo-Simulationen verwendet werden. Appendix B bietet eine Übersicht über wichtige R-Befehle.

Pflichtliteratur

Wooldridge (2009). *Introductory Econometrics. A Modern Approach*, 4. Auflage oder neuer, Thomson South-Western (Chapters 5, 7, 9, 10 - 12, 18, Appendix B, C)

Ergänzungsliteratur

Deutsch

- Deistler, M. und Scherrer, W. (2018). *Modelle der Zeitreihenanalyse*, Birkhäuser. (Im UR-Netz online verfügbar.)

- Kirchgässner, G. und Wolters, J. (2006). *Einführung in die moderne Zeitreihenanalyse*, Vahlen.

- Neusser, K. und Wagner, M. (2022). *Zeitreihenanalyse in den Wirtschaftswissenschaften*, 4. Auflage, Springer Gabler. (Im UR-Netz online verfügbar, Ausgabe (2006) zum Ausleihen verfügbar.)

  Schlittgen, R. und Sattarhoff, Christina (2020). *Angewandte Zeitreihenanalyse mit R*, De Gruyter.

Englisch:

- Brooks, C. (2008, 2019). *Introductory econometrics for finance*, 2nd ed., Cambridge University Press.
- Diebold, F.X. (2007). *Elements of forecasting*, 4. ed., Thomson/South-Western.
- Enders, W. (2015). *Applied econometric time series*, 4. ed., Wiley.
- Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University Press.
- Kirchgässner, G., Wolters, J. and Hassler, U. (2013). *Introduction to modern time series analysis*, 2nd. ed., Springer, Berlin.
- Lütkepohl, Helmut und Krätzig, Markus (2004). *Applied Time Se-*

ries Econometrics, Cambridge University Press. (Im UR-Netz online verfügbar; Version 2008 ist ein Reprint.)

1. Einleitung und Überblick

1.1. Beispiele von Regressionsmodellen mit Zeitreihen

1.1.1. Phillips-Kurve

Ursprüngliche Version der Phillips-Kurve nach Phillips (1958), der beobachtete: Es besteht ein Zusammenhang zwischen der (Lohn-) Inflations- und der Arbeitslosenrate.

- Grundlage für Wirtschaftspolitik?
- Grundlage für Prognosen?

Jährliche Arbeitslosen- und Inflationsraten für Deutschland in Abbildung 1.1

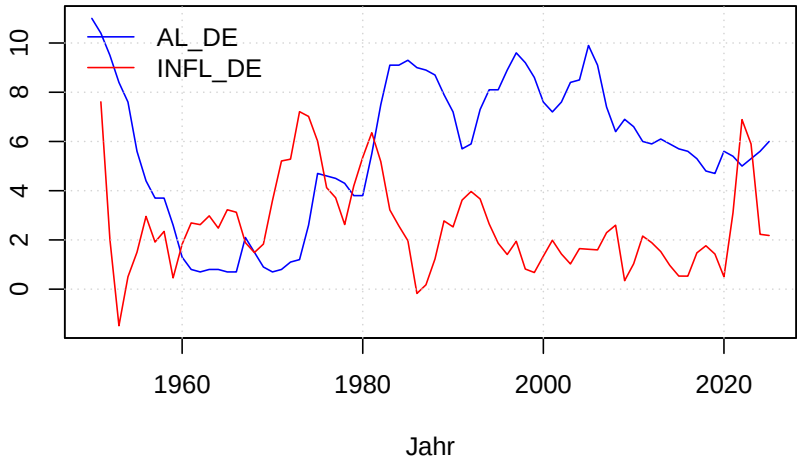

Abbildung 1.1.: Jährliche Arbeitslosenrate für Westdeutschland und Inflationsraten für Deutschland von 1950 bis 2025 (Quelle: Deutsche Bundesbank), siehe Appendix A.1 für das R-Programm

Mögliches Regressionsmodell:

$$inf_t = \beta_0 + \beta_1 alr_t + u_t, \quad t = 1, 2, \ldots$$

Beispiel eines **statischen Regressionsmodells**.

Geschätztes Modell auf Basis von Jahresdaten von 1950 bis 1970:

$$\underset{}{inf_t} = \underset{(0.32)}{2.77} - \underset{(0.08)}{0.26}\, alr_t + \hat{u}_t$$

- Interpretation: Sinkt die Arbeitslosenrate um einen Prozentpunkt, steigt die Inflation um 0.3 Prozentpunkte.
- korrekte Spezifikation?
- zuverlässige Schätzergebnisse?
- Zusammenhang über die Zeit hinweg stabil?

Alternative (ökonometrische) Spezifikationen:

- zusätzlich verzögerte exogene Regressoren

$$inf_t = \beta_0 + \beta_1 alr_t + \beta_2 alr_{t-1} + u_t, \quad t = 1, 2, \ldots$$

  Beispiel eines **Regressionsmodels mit (streng?) exogenen Regressoren**.

- zusätzlich verzögerte endogene Variablen als Regressoren

$$inf_t = \beta_0 + \beta_1 alr_t + \beta_2 inf_{t-1} + u_t, \quad t = 1, 2, \ldots$$

  Beispiel eines **dynamischen Regressionsmodells**.

- ausschließlich verzögerte endogene Variablen als Regressoren

$$inf_t = \beta_0 + \beta_1 inf_{t-1} + u_t, \quad t = 1, 2, \ldots$$

  Beispiel eines (univariaten) **Autoregressionsmodells**.

1.1.2. Prognose von Aktienindizes

Entwicklung des realen US-Aktienindex S&P 500 in Abbildung 1.2.

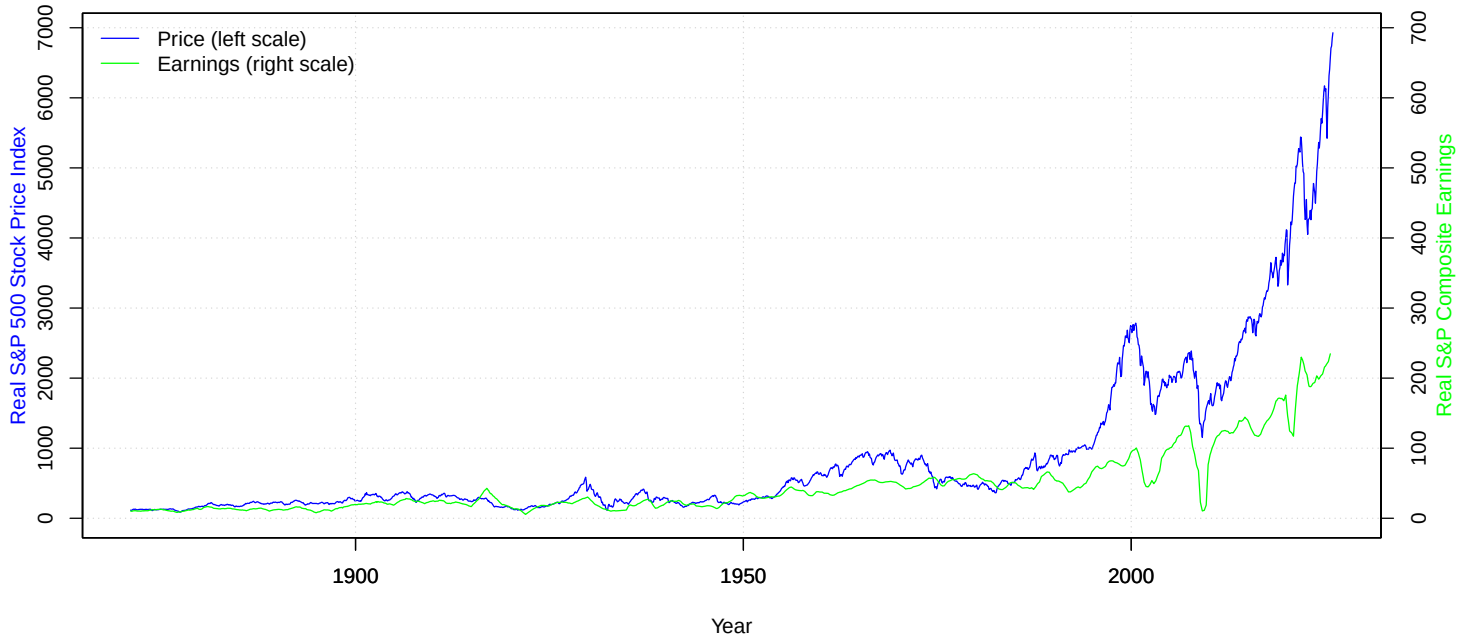

Abbildung 1.2.: Realpreise des S&P 500 Composite und reale Gewinne Januar 1871 - Januar 2026 (Quelle: Homepage von Robert Shiller: http://www.econ.yale.edu/~shiller/data.htm), siehe Appendix A.2 für das R-Programm

Mögliches autoregressives Modell:

$$P_t = \nu + \rho P_{t-1} + e_t, \quad t = 1, 2, \ldots$$

Geschätztes Modell:

$$P_t = \underset{(1.45)}{-2.00} + \underset{(0.001)}{1.007}\, P_{t-1} + \hat{e}_t, \quad \hat{\sigma} = 49,89$$

- korrekte Spezifikation?
- zuverlässige Schätzergebnisse, Standard t-Tests verwendbar?
- Zusammenhang über die Zeit hinweg stabil?
- Index prognostizierbar?

1.2. Was wird in Zeitreihenkonometrie behandelt?

- Schätz- und Testeigenschaften des OLS-, GLS-, FGLS-Schätzers, wenn die *Störterme nicht normalverteilt* sind.
- Ökonometrie für Zeitreihendaten:
  - Modelle mit *verzögerten exogenen* Regressoren,
  - Modelle mit *verzögerten endogenen* Regressoren,
  - *Autoregressive Modelle* der Ordnung p:

$$y_t = \nu + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + u_t,$$

  - Modellierung von *Trends* und *Saisonkomponenten* in Zeitreihendaten, vgl. Abbildungen 1.3 und 1.4.

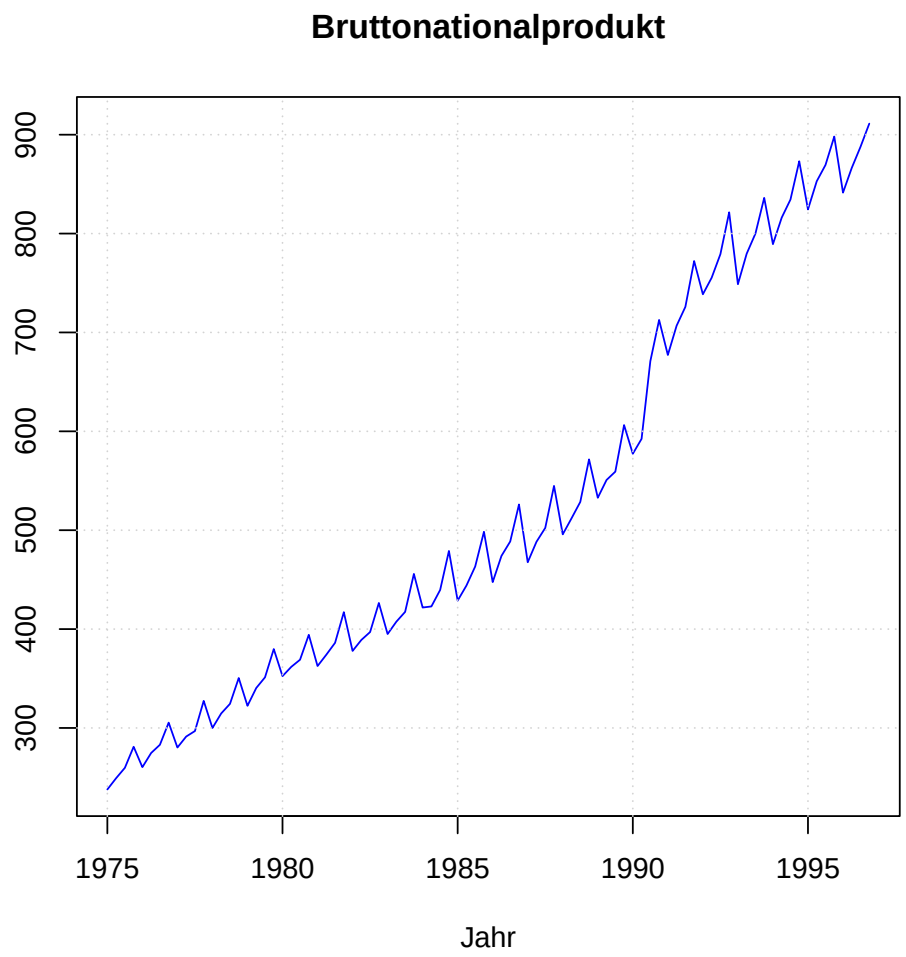

Abbildung 1.3.: Vierteljahresdaten des deutschen Bruttonationalprodukts, 1975 bis 1996 (Westdeutschland bis 1990:II, saisonal nichtbereinigt; Quelle: Beispieldatensatz GermanGNP.dat in JMulTi, Originalquelle: DIW), siehe Appendix A.3 für das R-Programm

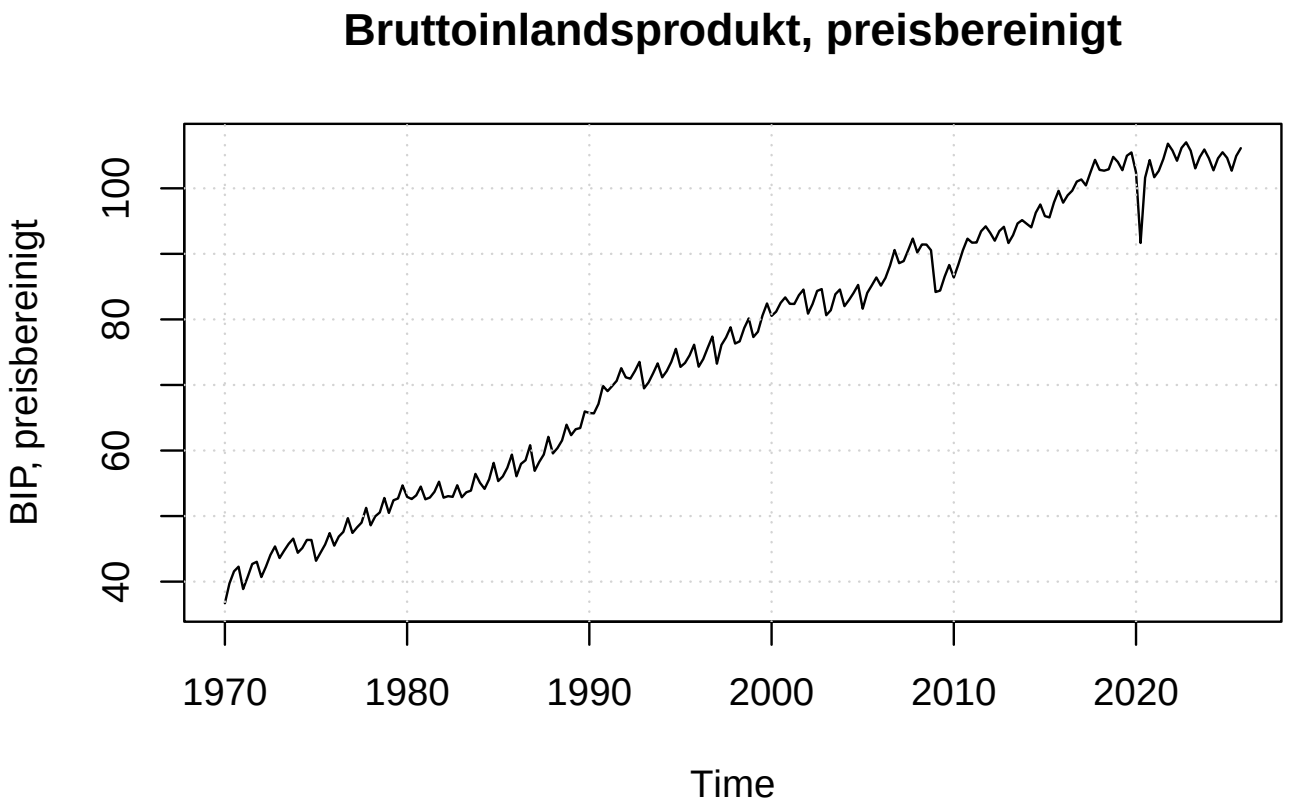

Abbildung 1.4.: Vierteljahresdaten des deutschen realen Bruttoinlandsprodukts 1970-2024 (Westdeutschland bis 1990:VI, Einheit 2015=100; Quelle: Deutsche Bundesbank, BBNZ1.Q.DE.N.H.0000.L), siehe Appendix A.3 für das R-Programm

– Regressionsmodelle mit heteroskedastischen und/oder autokorrelierten Fehlern, z.B.:

$$\begin{aligned} y_t &= \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + u_t, \\ u_t &= \nu + \alpha u_{t-1} + \varepsilon_t, \quad Var(\varepsilon_t) = \sigma_t^2. \end{aligned}$$

– *Vektorautoregressive Modelle*:
zur Modellierung von Interaktionen zwischen ökonomischen Variablen. Beispiel:

$$\begin{aligned} inf_t &= \nu_1 + a_{11}\, inf_{t-1} + a_{12}\, alr_{t-1} + u_t, \\ alr_t &= \nu_2 + a_{21}\, inf_{t-1} + a_{22}\, alr_{t-1} + v_t. \end{aligned}$$

– *Kointegrationsmodelle* zur Schätzung langfristiger Gleichgewichtsbeziehungen zwischen ökonomischen Variablen, z.B. zwischen Inflation und Arbeitslosenrate (falls diese existiert).

1.3. Voraussetzungen für Zeitreihenökonometrie aus Einführung in die Ökonometrie

- Einfaches lineares Regressionsmodell.
- Multiples lineares Regressionsmodell, OLS:
  - Spezifikation der Regressionsfunktion,
  - Interpretation der Parameter,
  - Schätzeigenschaften des OLS-Schätzers unter Standardannahmen (Erwartungswert, Varianz, Wahrscheinlichkeitsverteilung, Vergleich zu anderen Schätzern),
  - Schätzeigenschaften des OLS-Schätzers bei verletzten Annah-

men,

  - Technik: einfache Matrizenrechnung,
  - Modellselektion (Tests, Modellselektionskriterien, Bestimmtheitsmaß),
  - Tests (t-Tests, F-Tests) und Konfidenzintervalle,
  - Prognose.

- Multiples lineares Regressionsmodell, Erweiterungen:
  - Heteroskedastie (Tests),
  - GLS-Schätzer (Annahmen, Schätzeigenschaften), FGLS-Schätzer und OLS mit White Standardfehlern.

1.4. Zeitreihendaten und stochastische Prozesse

- **Zeitreihendaten** werden sequentiell über verschiedene Zeitpunkte/ -perioden erhoben, zum Indizieren wird gewöhnlich der Index t benutzt.
- Zeitreihendaten unterscheiden sich hinsichtlich ihrer Beobachtungshäufigkeit
  - täglich,
  - wöchentlich,
  - monatlich, etc.
  - vierteljährlich (Quartalsdaten)

– jährlich

- Sind Zeitreihenbeobachtungen Zufallsvariablen? Ja, z.B. ist morgiger DAX-Wert heute unbekannt. Ebenso: Die Zeitreihe der täglichen DAX-Werte im kommenden Jahr enthält unbekannte Werte. Diese Werte stellen eine *Folge* von Zufallsvariablen mit dem Index Zeit dar.

  Also:

  Definition: Eine Folge von Zufallsvariablen mit einem Zeitindex wird als **Zeitreihenprozess** oder als **stochastischer Prozess** bezeichnet.

- **Wichtiger Unterschiede** zu Querschnittsdaten:

  - Das mehrmalige Ziehen von Stichproben, wie es häufig bei Querschnittsdaten möglich ist, geht nicht. Man kann genau eine Stichprobe des täglichen Euro/Yen-Kurses von 2002 bis 2004 beobachten. **Man sagt**: Man kann *genau eine* **Realisation** des stochastischen Prozesses beobachten. Man kann nicht die Geschichte zurückdrehen und den ökonomischen Verlauf nochmals beobachten. Aber a priori war dieser Verlauf nicht bekannt, da vor 2002 eine Vielzahl *anderer* Verläufe denkbar waren.

  - Wie man später sehen wird, ist die Annahme MLR.2 "Die (y_t, x_t) sind rein zufällig gezogen" sehr häufig verletzt — siehe später.

- Möchte man die **dynamische Wirkung** von Einflussfaktoren bzw. Störgrößen auf die abhängige Variable untersuchen, muss man Zeitreihendaten verwenden.

- Oftmals liegen für bestimmte Fragestellungen nur Zeitreihendaten vor.

- Zur Vereinfachung der Schreibweise werden alle Regressoren zum Zeitpunkt t in einem Zeilenvektor zusammengefasst:

$$\mathbf{x}_t = \begin{pmatrix} x_{t1} & x_{t2} & \cdots & x_{tk} \end{pmatrix}.$$

- Beachte: Sind beispielsweise x_{t1} bis x_{tk} jeweils einzelne stochastische Prozesse, lassen sich die einzelnen Zeitreihenprozesse in einem gemeinsamen Prozess $(y_t, \mathbf{x}_t)$, $t = 1, 2, \ldots, T$ zusammenfassen. Da $(y_t, \mathbf{x}_t)$ als Vektor geschrieben werden kann, sagt man auch, dass $\{(y_1, \mathbf{x}_1), (y_2, \mathbf{x}_2), \ldots, \ldots, (y_T, \mathbf{x}_T)\}$ ein Vektorzeitreihenprozess ist.

2. Regressionsmodelle für Zeitreihendaten I: Regressionsmodelle mit streng exogenen Regressoren

Beispiel aus der Marktetingforschung

- Auf welche Weise hängt die Nachfrage nach Bier von den Werbeausgaben ab?

- Verfügbare Daten (für die Niederlande):
  zweimonatliche Daten von 1978 bis 1984 (Datenquelle: Franses (1991). Primary Demand for Beer in the Netherlands: An Application of the ARMAX Model Specification, *Journal of Marketing Research* 28, p. 240-5):
  - q: Endnachfrage für Bier (Liter/Bevölkerung über 15),
  - wa: Gesamtausgaben für Werbung (cent (niederländische Gulden) / Bevölkerung über 15),
  - $temp$: durchschnittliche Tagestemperatur (Grad Celcius).
- Einfachstes (statisches) Regressionsmodell:

$$q_t = \beta_0 + \beta_1 wa_t + \beta_2 temp_t + u_t.$$

2.1. Statische und Finite-Distributed-Lag-Modelle

2.1.1. Statische Modelle

- Ein Zeitreihenmodell heißt **statisch**, wenn nur Variablen mit dem gleichen Zeitindex in das Modell eingehen:

$$y_t = \beta_0 + \beta_1 z_t + \beta_2 x_t + u_t.$$

  - Beispiel: Einfache Phillipskurve:

$$inf_t = \beta_0 + \beta_1 \, alr_t + u_t, \quad t = 1, 2, \ldots.$$

  - Beispiel: Biernachfragegleichung:

$$q_t = \beta_0 + \beta_1 \, wa_t + \beta_2 \, temp_t + u_t.$$

- Liegt Autokorrelation in den Residuen vor (Testmethoden später), ist das im Allgemeinen ein Hinweis darauf, dass ein statisches Modell nicht geeignet ist.

- Auch sind statistische Modelle per Konstruktion ungeeignet, dynamische Effekte zu analysieren.

2.1.2. Finite-Distributed-Lag-Modelle

- Ein Modell wird als **Finite-Distributed-Lag-(FDL) Modell** bezeichnet, wenn es **verzögerte exogene Variablen** enthält:

$$y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \cdots + \delta_q z_{t-q} + \beta_1 x_t + u_t.$$

Beispiel: Biernachfragegleichung:

$$q_t = \beta_0 + \beta_1\, wa_t + \beta_2\, wa_{t-1} + \beta_3\, wa_{t-2} + \beta_4\, temp_t + u_t.$$

- Diese Modelle können mit KQ geschätzt werden. Sind die z_t's und x_t's *streng exogen*, gelten die bekannten Schätzeigenschaften des KQ-Schätzers, siehe Abschnitt 2.2.

- Ist dies nicht der Fall, lassen sich die Schätzeigenschaften des KQ-Schätzers nur asymptotisch bestimmen, siehe Kapitel 5.

- **Interpretation** der Parameter:

  - Der Parameter δ_0 gibt den **kurzfristigen Multiplikator** (impact propensity, impact multiplier) an, denn wird z_t ceteribus paribus unerwartet um c erhöht, steigt y_t im Erwartungswert um $\delta_0 c$ an, bzw.:

$$\frac{\partial E[y_t|z_t, z_{t-1}, \ldots, z_{t-q}, x_t]}{\partial z_t} = \delta_0.$$

  - Der **langfristige Multiplikator** (long-run propensity, long-run

multiplier) ist gegeben durch den Effekt, der nach q Perioden erreicht wird, wenn z dauerhaft um c erhöht wird:

$$E[y_t|z_t + c, z_{t-1} + c, \ldots, z_{t-q} + c, x_t] - E[y_t|z_t, z_{t-1}, \ldots, z_{t-q}, x_t] = (\delta_0 + \delta_1 + \cdots + \delta_q)c.$$

– Betrachtet man die Verteilung ("distribution") der δ_j, $j = 0, \ldots, q$, kann man etwas über die Dynamik einer dauerhaften Erhöhung um c auf y sagen.

- Welche Annahmen müssen wir treffen, um bei der KQ-Schätzung von statischen oder FDL-Regressionsmodellen die klassischen Schätzeigenschaften zu erhalten?

2.2. Wann gelten die klassischen Eigenschaften des KQ-Schätzers bei Zeitreihendaten?

Notation:

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} x_{10} & x_{11} & \cdots & x_{1k} \\ x_{20} & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{T0} & x_{T1} & \cdots & x_{Tk} \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$$

Für folgendes Regressionsmodell:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u},$$

lautet der KQ-Schätzer $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$.

Klassische Eigenschaften des KQ-Schätzers unter geeigneten Annahmen:

- Erwartungstreue,
- BLUE (best linear unbiased estimator),
- Testverteilungen exakt oder asymptotisch bekannt.

2.2.1. Annahmen für Erwartungstreue des KQ-Schätzers bei Zeitreihen

Benötigte Annahmen bei Zeitreihen:

- **Annahme TS.1** (entspricht praktisch MLR.1):

Der (Vektor-)Zeitreihenprozess $(y_t, \mathbf{x}_t)$, $t = 1, 2, \ldots, T$ ist linear in den Parametern:

$$y_t = \beta_0 + x_{t1}\beta_1 + \cdots + x_{tk}\beta_k + u_t.$$

- **Annahme TS.2 (Keine perfekte Kollinearität)** (entspricht MLR.3).

- **Annahme TS.3 (Bedingter Erwartungswert Null)**:

$$E[\mathbf{u}|\mathbf{X}] = 0 \quad \text{bzw.} \quad E[u_t|\mathbf{X}] = 0 \quad \text{für alle } t = 1, 2, \ldots, T.$$

Wiederholung (Erwartungstreue des KQ-Schätzers):

$$E[\hat{\boldsymbol{\beta}}|X] \overset{(TS.2)}{=} E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}|\mathbf{X}]$$

$$\overset{(TS.1)}{=} E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \mathbf{u})|\mathbf{X}] = \boldsymbol{\beta} + E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}|\mathbf{X}]$$

$$= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E[\mathbf{u}|\mathbf{X}] \overset{(TS.3)}{=} \boldsymbol{\beta}.$$

Bemerkungen:

- **TS.3 verschärft MLR.4**. MLR.4 fordert lediglich, dass der Fehler u_t von allen **gegenwärtigen** Regressoren unabhängig ist, also

$$E[u_t|\mathbf{x}_t] = 0.$$

  Da jedoch MLR.2 (Zufallsstichprobe) nicht mehr gefordert wird, ist für Schritt (2) weiter oben nunmehr TS.3 nötig.

- Variable, die TS.3 erfüllen, werden **streng exogen** genannt.

- Die Annahme strenger Exogenität ist häufig verletzt. Bspw. sind Rückwirkungen der abhängigen Variablen auf die Regressoren ausgeschlossen (Werbeausgaben dürfen nicht auf vergangenen Bierkonsum reagieren). Ausgeschlossen ist auch, dass der vergangene Bierkonsum auf den gegenwärtigen Bierkonsum wirkt!

2.2.2. Annahmen für Gültigkeit des Gauss-Markov-Theorems

- **Annahme TS.4 (Bedingte Homoskedastie)**:

$$Var(u_t|\mathbf{X}) = \sigma^2 \qquad \text{für alle } t = 1, 2, \ldots, T, \quad \overset{TS.3}{\Rightarrow} Var(u_t) = \sigma^2$$

Bemerkungen:

– Annahme TS.4 ist stärker als Annahme MLR.5, da die Regressoren **streng exogen** sein müssen (Grund: Annahme der Zufallsstichprobe unsinnig bei Zeitreihen).

– Ist Annahme TS.4 verletzt, aber die Art der Heteroskedastie bekannt, kann man wie im Fall von Querschnittsdaten den GLS-Schätzer bzw. den FGLS-Schätzer verwenden.

- **Annahme TS.5 (Keine (Auto)korrelation in den Störtermen)**:

$$Cov(u_t, u_s|\mathbf{X}) = 0 \quad \text{für alle } t \neq s,$$

bzw.

$$Corr(u_t, u_s|\mathbf{X}) = \frac{Cov(u_t, u_s|\mathbf{X})}{\sqrt{Var(u_t|\mathbf{X})Var(u_s|\mathbf{X})}} = 0 \quad \text{für alle } t \neq s.$$

Bemerkungen:

– Sind die Regressoren $\mathbf{X}$ und die Störterme $\mathbf{u}$ stoschastisch unabhängig, dann gilt:

$$Cov(u_t, u_s|\mathbf{X}) = Cov(u_t, u_s).$$

– Der Zusatz *Auto* verdeutlicht, dass bei Zeitreihen immer eine Richtung vorliegt. So beeinflusst u_s gegeben $t > s$ u.U. den weiter in der Zukunft liegenden Störterm u_t.

- Zur Diagnose und Lösungsansätzen von Autokorrelation in den Residuen siehe 8.1 bis 8.5.

- Annahme TS.5 fordert nichts über die Korrelation in den unabhängigen Variablen!

- **Varianz des Schätzers**: Gelten Annahmen TS.1 bis TS.5, lässt sich zeigen (siehe **Einführung in die Ökonometrie**):

$$Var(\hat{\boldsymbol{\beta}}|\mathbf{X}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1},$$

bzw. für einzelne $\hat{\beta}_j$:

$$Var(\hat{\beta}_j|\mathbf{X}) = \frac{\sigma^2}{\mathsf{SST}_j(1-R_j^2)}, \quad j = 1, 2, \ldots, k,$$

mit $SST_j := \sum_{t=1}^{T}\left(x_{tj} - \bar{x}_j\right)^2$ und $\bar{x}_j = T^{-1}\sum_{t=1}^{T} x_{tj}$. Das Bestimmtheitsmaß R_j^2 erhält man aus der Regression von x_j auf

alle anderen Spalten in $\mathbf{X}$.

- **Geschätzte Varianz**: Gelten Annahmen TS.1 bis TS.5, so ist $\hat{\sigma}^2$ ein unverzerrter Schätzer für σ^2:

$$\hat{\sigma}^2 = \frac{\text{SSR}}{T-k-1}.$$

- **Gauss-Markov-Theorem**: Gelten Annahmen TS.1 bis TS.5, gilt auch das Gauss-Markov-Theorem:
  Der KQ-Schätzer ist der *beste lineare unverzerrte Schätzer* gegeben $\mathbf{X}$.

  Linearer Schätzer: $\tilde{\boldsymbol{\beta}} = \mathbf{A}\mathbf{y}$, wobei $\mathbf{y}$ ausschließlich von $\mathbf{X}$ abhängen darf. Beispiel: KQ-Schätzer mit $\mathbf{A} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.

2.2.3. Tests

Zum **Testen** benötigt man noch eine Verteilungsannahme für die Störterme. Alle bisherigen exakten Ergebnisse für den KQ-Schätzer für **endliche** Stichprobengrößen bleiben gültig, wenn zusätzlich zu Annahmen TS.1 bis TS.5 gilt:

Annahme TS.6 (Normalverteilte Fehler):

$$\mathbf{u}|\mathbf{X} \sim N(\mathbf{0}, \sigma^2\mathbf{I}).$$

D.h., die Störterme u_t sind gegeben $\mathbf{X}$ unabhängig und identisch gemäß $N(0, \sigma^2)$ verteilt.

Bemerkung:

- TS.6 impliziert TS.3 (Strenge Exogenität), TS.4 (Homoskedastie) und TS.5 (Keine Autokorrelation).

2.2.4. Zusammenfassung

- Unter Annahmen TS.1, TS.2 und TS.6 weist der KQ-Schätzer die bekannten schönen Eigenschaften auf, d.h. die Schätzeigenschaften gelten exakt für endliche Stichproben, also für kurze wie lange Zeitreihen. Damit lassen sich, wie bisher (Ausnahme FGLS-Schätzer), alle t-, und F-Tests durchführen.
  (Hinweis: in Wooldridge (2009) etwas andere Def. TS.6, deshalb dort TS.1 bis TS.6 notwendig.)

- Die Annahme *strenger Exogenität* ist in der Praxis häufig verletzt. (Beispiel: Die Werbeausgaben für Bier reagieren vermutlich auf den vergangenen Bierkonsum.)

- Ist die Annahme strenger Exogenität verletzt, lässt sich diese abschwächen. Allerdings gelten dann die Schätzeigenschaften des KQ-Schätzers nur noch **asymptotisch** (siehe 5.3). Mit anderen Worten: Man kennt die Schätzeigenschaften umso besser, je größer die Stichprobe ist, d.h. je mehr Zeitreihenbeobachtungen vorliegen.

2.3. Beispiel: Schätzung der Biernachfrage

2.3.1. Statisches Regressionsmodell

Geschätztes statisches Regressionsmodell in Tabelle 2.1.

Bemerkungen zum geschätzten Modell:

- Sind die Regressoren wirklich streng exogen?
- Gibt es Autokorrelation in den Residuen?
  Man beachte, dass hier der Zeitindex eine natürliche Anordnung der Residuen erlaubt (im Gegensatz zu Querschnittsdaten). Abbildung 2.1 zeigt die Residuen.
- Haben verzögerte exogene Variable einen Erklärungswert?

Tabelle 2.1.: Biernachfrage, KQ

```
Call:
lm(formula = q ~ wa + temp, data = beer_dat)

Residuals:
    Min      1Q  Median      3Q     Max
-5.2900 -1.1754  0.0456  0.8566  6.5502

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 15.7283406  1.8074644   8.702 1.12e-10 ***
wa          -0.0001433  0.0007435  -0.193    0.848
temp         0.2645286  0.0559040   4.732 2.91e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.938 on 39 degrees of freedom
Multiple R-squared:  0.3667,Adjusted R-squared:  0.3342
F-statistic: 11.29 on 2 and 39 DF,  p-value: 0.0001352

> SelectCritEViews(ols_beer_1)
     aic      hq      sc
[1,] 4.229486 4.27498 4.353605
```

Anmerkungen: Daten verfügbar unter: beer_dat.txt. Quelle: Franses (1991), siehe Appendix A.4 für das R-Programm.

Abbildung 2.1 zeigt die Struktur der Residuen.

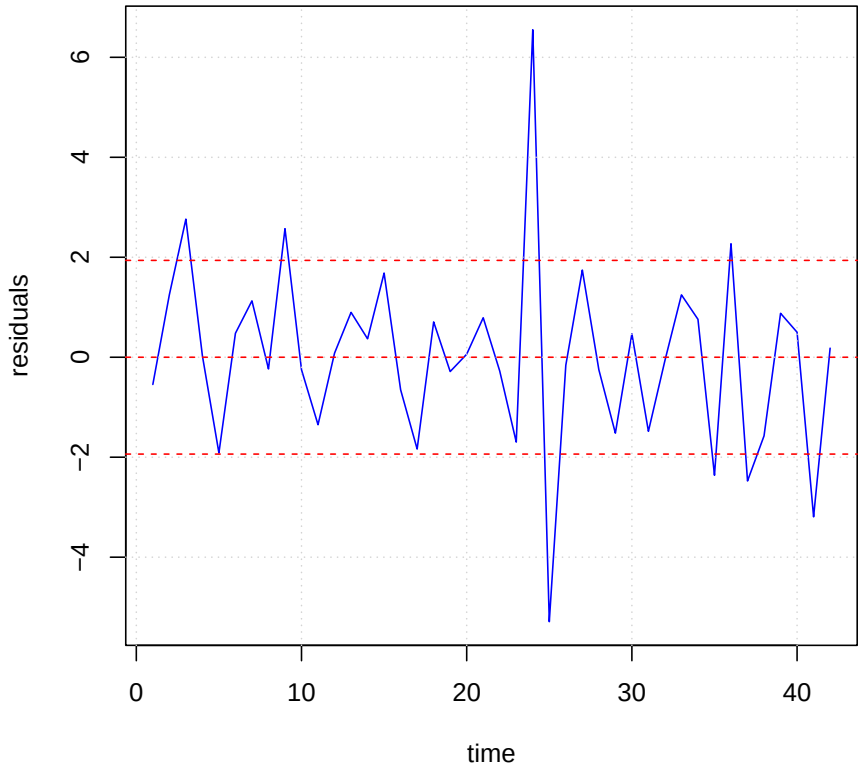

Abbildung 2.1.: Residuen.

Anmerkungen: Residuen $\hat{\mathbf{u}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$.

2.3.2. FDL-Modell

Geschätztes FDL-Regressionsmodell in Tabelle 2.2.

Interpretation zu Tabelle 2.2:

- Würde man bei dieser Spezifikation bleiben und werden die Werbeausgaben um 100 Einheiten erhöht, ist der kurzfristige Multiplikatoreffekt 0.01 (statistisch insignifikant), der langfristige Multiplikatoreffekt hingegen

$$(0.0001 + 0.0015 - 0.0024) \cdot 100 = -0.0008 \cdot 100 = -0.08,$$

  also negativ!

Tabelle 2.2.: Biernachfrage, FDL

```
Time series regression with "ts" data:
Start = 1978(3), End = 1984(6)

Call:
dynlm(formula = q ~ 1 + L(wa, 0:2) + temp, data = beer_dat)

Residuals:
    Min      1Q  Median      3Q     Max
-3.9332 -1.0346 -0.0160  0.9204  4.9738

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 17.1363341  2.4419816   7.017 3.63e-08 ***
L(wa, 0:2)0  0.0001036  0.0006880   0.151  0.88118
L(wa, 0:2)1  0.0015218  0.0007146   2.130  0.04031 *
L(wa, 0:2)2 -0.0023953  0.0007092  -3.377  0.00181 **
temp         0.2654756  0.0509202   5.214 8.41e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.703 on 35 degrees of freedom
Multiple R-squared:  0.5486,Adjusted R-squared:  0.497
F-statistic: 10.64 on 4 and 35 DF,  p-value: 9.545e-06

     aic      hq       sc
[1,] 4.019561 4.095892 4.230671
```

Anmerkungen: Daten: `beer_dat.txt`. Quelle: Franses (1991), siehe Appendix A.4 für das R-Programm.

2.3.3. Autokorrelation

Ergibt die Residuenanalyse, dass Autokorrelation in den Residuen vorliegt, kann man es mit weiteren verzögerten exogenen Variablen probieren. Dies hilft jedoch nicht, wenn die verzögerte endogene Variable — im Beispiel des Bierkonsums in den 2 vorhergehenden Monaten — einen Einfluss hat. Man muss dann ein **dynamisches Regressionsmodell** schätzen:

$$q_t = \beta_0 + \beta_1 q_{t-1} + \beta_2 wa_t + \beta_3 wa_{t-1} + \beta_{t-2} wa_{t-2} + \beta_5 temp_t + u_t.$$

Um die Eigenschaften des KQ-Schätzers bestimmen zu können, müssen die Annahmen TS.1 bis TS.6 abgeschwächt werden $\longrightarrow$ Kapitel 7 der Veranstaltung.

Zu lesen: Chapter 10.1 - 10.3 in Wooldridge (2009).

2.4. Wahl von Datentransformationen und funktionaler Form

Vorbemerkung: Funktionale Form versus Datentransformation

- Wird die Zeitreihe eines Regressors vor der Modellbildung logarithmiert, z.B. $x = \log(w)$, impliziert dies gleichzeitig eine Veränderung der funktionalen Form bezüglich des Zusammenhangs zwischen y und w.

  Beispiel einer Einfachregression:
  level-level Modell

$$y_t = \beta_0 + \beta_1 w_t + u_t, \quad t = 1, 2, \ldots, T$$

versus

level-log Modell

$$y_t = \beta_0 + \beta_1 \log(w_t) + u_t, \quad t = 1, 2, \ldots, T.$$

Der bedingte Erwartungswert $E[y|w]$ lautet jeweils:

$$E[y|w] = \begin{cases} \beta_0 + \beta_1 w & \text{level-level Modell,} \\ \beta_0 + \beta_1 \log(w) & \text{level-log Modell.} \end{cases}$$

- Eine Datentransformation wird häufig durchgeführt, um eine gewünschte Interpretation der Modellparameter zu ermöglichen (z.B. als Elastizitäten). Beachte: Dies geht nur, solange die Modellannahmen nicht verletzt werden.

- Eine Datentransformation ist immer sinnvoll, wenn dadurch die Modellannahmen besser erfüllt werden können.

2.4.1. Lineare Regressionsmodelle und nichtlineare funktionale Form

- Annahme TS.1 fordert, dass

$$y_t = \beta_0 + \beta_1 x_{t1} + \cdots + \beta_k x_{tk} + u_t.$$

  Sie besagt, dass das Regressionsmodell *linear in den Parametern* $\beta_0, \ldots, \beta_k$ ist, nicht jedoch in den Variablen selbst!

- Bezüglich einer mit w_t bezeichneten Ausgangsvariablen sind viele Datentransformationen und damit verschiedene funktionale Formen zwischen y und w möglich. Einige Beispiele:

  – logarithmierte Variable: $x_{t1} = \log(w_t)$,

  – quadratische Terme: $x_{t1} = w_t$, $x_{t2} = w_t^2$,

  - Quadrate von Logarithmen: $x_{t1} = \log(w_t)$, $x_{t2} = (\log(w_t))^2$,
  - Interaktionsterme: $x_{t1} = w_t$, $x_{t2} = w_t z_t$,
  - und viele mehr.

- Man beachte, dass sich entsprechend die Interpretation der Parameter ändert.
- Derartige Regressionsmodelle dienen häufig dazu, nichtlineare funktionale Formen, deren Regressionsfunktion nichtlinear in den Parametern sind – und damit schwerer zu schätzen sind –, zu approximieren.

2.4.2. Modelle mit logarithmierten Variablen

- In statischen Modellen lassen sich log-log, log-level, level-log Modelle wie bei Querschnittsdaten interpretieren.
- In FDL-Modellen in log-log Form

  $$\log(y_t) = \alpha_0 + \delta_0 \log(z_t) + \delta_1 \log(z_{t-1}) + \cdots + \delta_q \log(z_{t-q}) + \beta_1 \log(x_t) + u_t$$

  gilt:

level-level	log-log
kurzfristiger Multiplikator	kurzfristige Elastizität
langfristiger Multiplikator	langfristige Elastizität

Beispiel: log-log FDL-Modell der Biernachfrage

Spezifikation:

$$\log(q_t) = \alpha_0 + \delta_0 \log(wa_t) + \delta_1 \log(wa_{t-1}) + \delta_2 \log(wa_{t-2}) + \\ + \beta_1 \log(temp_t) + u_t.$$

Interpretation von log-log Modell und Schätzergebnissen (Tabelle 2.3):

- kurzfristige Elastizität (short-run elasticity): $\hat{\delta}_0 = 0.0068$,
- langfristige Elastizität (long-run elasticity): $\hat{\delta}_0 + \hat{\delta}_1 + \hat{\delta}_2 = -0.0923$. Die permanente Erhöhung der Werbeausgaben um 1% führt zu einer Senkung der Biernachfrage (nach 2 Doppelmonaten) um 0.09%.
- Frage: Wie lässt sich überprüfen, ob die langfristige Elastizität signifikant ist? Voraussetzungen hierfür?

Tabelle 2.3.: Biernachfrage, log-log FDL-Modell

```
Time series regression with "zoo" data:
Start = 1978(3), End = 1984(6)

Call:
dynlm(formula = log(q) ~ L(log(wa), 0:2) + log(temp), data = beer_dat_zoo)

Residuals:
      Min        1Q    Median        3Q       Max
-0.257547 -0.067136 -0.004547  0.065998  0.276699

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       3.323038   1.154452   2.878  0.00686 **
L(log(wa), 0:2)1  0.006769   0.103839   0.065  0.94841
L(log(wa), 0:2)2  0.225588   0.101310   2.227  0.03270 *
L(log(wa), 0:2)3 -0.324634   0.100521  -3.230  0.00275 **
log(temp)         0.127032   0.024844   5.113 1.22e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1008 on 34 degrees of freedom
  (1 Beobachtung als fehlend gelöscht)
Multiple R-squared:  0.5545,Adjusted R-squared:  0.502
F-statistic: 10.58 on 4 and 34 DF,  p-value: 1.12e-05

> SelectCritEViews(ols_beer_3b)
     aic       hq        sc
[1,] -1.632925 -1.556403 -1.419648
```

Anmerkungen: Daten: beer_dat.txt. Quelle: Franses (1991), siehe Appendix A.4 für das R-Programm.

Log-(level) Modelle: Große Veränderungen in einem Level-Regressor

Es sei folgendes Regressionsmodell gegeben:

$$\log(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u, \tag{2.1}$$

wobei x_2 eine Level-Variable ist.

Wie lässt sich der Einfluss von x_2 *exakt* bestimmen bzw. wie lautet die Interpretation von β_2? Aus (2.1) folgt:

$$y = e^{\log(y)} = e^{\beta_0+\beta_1 x_1+\beta_2 x_2+u} = e^{\beta_0+\beta_1 x_1+\beta_2 x_2} \cdot e^u$$

und für den konditionalen Erwartungswert:

$$E(y|x_1, x_2) = e^{\beta_0+\beta_1 x_1+\beta_2 x_2} \cdot E(e^u|x_1, x_2). \tag{2.2}$$

Einsetzen von zwei Ausprägungen der Level-Variablen x_2 in (2.2) ergibt:

$$\begin{aligned} E(y|x_1, x_2) &= e^{\beta_0+\beta_1 x_1+\beta_2 x_2} \cdot E(e^u|x_1, x_2), \\ E(y|x_1, x_2 + \Delta x_2) &= e^{\beta_0+\beta_1 x_1+\beta_2 x_2} \cdot E(e^u|x_1, x_2 + \Delta x_2) \cdot e^{\beta_2 \Delta x_2} \\ &= E(y|x_1, x_2) \cdot e^{\beta_2 \Delta x_2}, \end{aligned}$$

wobei vorausgesetzt wurde, dass $E(e^u|x_1, x_2 + \Delta x_2)$ bezüglich Δx_2 konstant ist (z.B., wenn u von x_1, x_2 stochastisch unabhängig ist). Daraus ergibt sich die relative Änderung im durchschnittlichen Wert der abhängigen Variable mit:

$$\begin{aligned} \frac{\Delta E(y|x_1, x_2)}{E(y|x_1, x_2)} &= \frac{E(y|x_1, x_2 + \Delta x_2) - E(y|x_1, x_2)}{E(y|x_1, x_2)} \\ &= \frac{E(y|x_1, x_2) \cdot e^{\beta_2 \Delta x_2} - E(y|x_1, x_2)}{E(y|x_1, x_2)} \\ &= e^{\beta_2 \Delta x_2} - 1. \end{aligned}$$

Damit erhalten wir:

$$\%\Delta E(y|x_1, x_2) = 100 \left(e^{\beta_2 \Delta x_2} - 1 \right).$$

Allgemein formuliert für den Fall von k Regressoren gilt:

$$\%\Delta E(y|x_1, x_2, \ldots, x_j, \ldots, x_k) = 100 \left(e^{\beta_j \Delta x_j} - 1 \right). \qquad (2.3)$$

Gleichung (2.3) gibt den exakten Partialeffekt an, während die Interpretation als Semi-Elastizität manchmal nur eine grobe Annäherung bietet.

Die exakte Berechnung des exakten Partialeffekts ist häufig sehr wichtig bei Dummyvariablen.

Beispiel: Weiterbildung und Ausschussquote

Untersucht werden soll, inwieweit die Ausschussrate von 157 Produktionsfirmen von Weiterbildungsmaßnahmen für Firmenangestellte abhängen (Daten: `jtrain` in R-Paket `wooldridge`).

Die Variable *scrap* gibt die Zahl der defekten Teile pro 100 an. Die Variable *hrsemp* gibt die jährliche pro-Kopf Weiterbildungszeit in Stunden einer Firma an (*sales*: U.S. Dollar Jahresumsatz; *employ*: Zahl der durchschnittlich Beschäftigten). Alle Zahlen beziehen sich auf das Jahr 1987.

Spezifikation:

$$\log(scrap) = \beta_0 + \beta_1 hrsemp + \beta_2 \log(sales) + \beta_3 \log(employ) + u.$$

Schätzergebnis: Die KQ-Schätzung liefert $\hat{\beta}_1 = -0.042$. Siehe Appendix A.5 für das R-Programm.

- "Standard"-Berechnung und -Interpretation:
  Erhöht man bspw. die Weiterbildungszeit um 10 h/Jahr, verringert sich die Ausschußrate um 42% (im ø, c.p.).

- Exakte Berechnung:
  Einsetzen in (2.3) ergibt hingegen $100(e^{-0.042 \cdot 10} - 1) = -34.3$, d.h. die Ausschußrate verringert sich im o.a. Fall um ca. 34%!

2.4.3. Modelle mit quadrierten Regressoren

Beispiel:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + u.$$

Marginaler Effekt einer Änderung in x_2 ist wegen

$$\frac{\partial E[y|x_1, x_2]}{\partial x_2} = \beta_2 + 2\beta_3 x_2$$

gegeben durch

$$(\beta_2 + 2\beta_3 x_2)\Delta x_2,$$

d.h. dessen Höhe hängt auch vom Niveau von x_2 ab (und die Interpretation von β_2 alleine macht wenig Sinn!).

Beispiel: Regensburger Mietspiegel von 1996

Geschätzte Gleichung :

$$NM = 356.64 + 6.19 \cdot WFL - 13.26 \cdot ALTER + 0.11 \cdot ALTER^2 + \hat{u}.$$

Die Änderung in der Nettomiete NM bei einem zusätzlichen Alters-jahr der Wohnung unterscheidet sich bspw. für eine Wohnung mit $ALTER = 10$ von einer Wohnung mit $ALTER = 100$ beträchtlich.

Anmerkungen:

- Verwendung von Querschnittsdaten.
- Obige Schätzgleichung entspricht nicht dem im Mietspiegel 1996 verwendeten Modell.
- Daten nicht für Unterrichtszwecke verfügbar.

Die zugehörigen marginalen Effekte ergeben sich mit

$$(-13.26 + 2 \cdot 0.11 \cdot 10) = -11.06$$

bzw.

$$(-13.26 + 2 \cdot 0.11 \cdot 100) = 8.74.$$

Der marginale Effekt ist gerade Null, wenn $\beta_2 + 2\beta_3 x_2 = 0$, also wenn $x_2 = -\beta_2/(2\beta_3)$. Es ist oft hilfreich, das Extremum zu bestimmen.
In unserem Mietspiegelbeispiel ergibt sich ein Minimum bei:

$$ALTER^* = 13.26/2 \cdot 0.11 = 60.3,$$

(was plausibel ist, da die Regensburger 30er Jahre Bauten oft eine besonders schlechte Bausubstanz aufweisen).

2.4.4. Quadrate von logarithmierten Variablen

Beispiel: Um nicht-konstante Elastizitäten zu approximieren, könnte man bspw. wie folgt ansetzen:

$$\log(y) = \beta_0 + \beta_1 x_1 + \beta_2 \log(x_2) + \beta_3 (\log(x_2))^2 + u.$$

Die Elastizität von y bzgl. x_2 lautet (überprüfen!):

$$\beta_2 + 2\beta_3 \log(x_2)$$

und ist nur dann konstant, falls $\beta_3 = 0$ gilt (siehe Section 6.2 in Wooldridge (2009)).

2.4.5. Interaktionsterme

Beispiel:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2 x_1 + u.$$

Marginaler Effekt einer Änderung in x_2 ist gegeben durch:

$$\Delta E(y|x_1, x_2) = (\beta_2 + \beta_3 x_1)\Delta x_2,$$

d.h. jetzt hängt der marginale Effekt auch von x_1 ab!

2.4.6. Wahl zwischen nicht-genesteten Modellen

Definition: *Nicht-genestet* bedeutet, dass sich ein Modell nicht als Spezialfall des anderen (und umgekehrt) darstellen lässt.

Beispiel:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + u$$

und

$$y = \beta_0 + \beta_1 \log(x_1) + u.$$

Eine Wahl zwischen beiden nicht-genesteten Modellen kann mit SC, AIC bzw. HQ (oder $\bar{R}^2$) erfolgen. Warum?

Zu lesen: Sections 6.1, 6.2, 6.3 und 10.4 in Wooldridge (2009).

2.5. Dummyvariable: Wiederholung und Vertiefung

2.5.1. Dummyvariable oder binäre Variable

Eine binäre Variable kann exakt zwei verschiedene Werte annehmen und ermöglicht so, zwei qualitativ unterschiedliche Zustände zu beschreiben. Beispiele: männlich vs. weiblich, arbeitslos vs. nichtarbeitslos, etc.

- Am besten wählt man für die beiden Werte 0 und 1, da sich dann die Koeffizienten von Dummyvariablen gut interpretieren lassen.
- Beispiel für eine Regressionsgleichung:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{k-1} x_{k-1} + \delta D + u,$$

wobei D entweder den Wert 0 oder 1 annimmt.

- Interpretation der Koeffizienten von Dummy-Variablen:

$$\begin{aligned} & E(y|x_1, x_2, D = 1) - E(y|x_1, x_2, D = 0) \\ & \quad = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{k-1} x_{k-1} + \delta \\ & \qquad - (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{k-1} x_{k-1}) \\ & \quad = \delta. \end{aligned}$$

Der Koeffizient einer Dummyvariable (oder "0-1-Variable") gibt die Verschiebung des Achsenabschnitts um δ bei Vorliegen $D = 1$ an. *Alle* Steigungsparameter β_i, $i = 1, \ldots, k-1$ bleiben *unverändert*.

Beachte: Bei der Interpretation von Dummyvariablen ist es wichtig, die Referenzgruppe zu kennen (d.h. die Periode bzw. Gruppe, für die die Dummyvariable 0 ist).

Beachte: Bei log-? Modellen ist hinsichtlich Dummyvariablen am besten der exakte Partialeffekt zu berechnen.

Beispiel: Verlagerung der Phillips-Kurve

Hat sich die Phillipskurve über die Jahrzehnte verschoben?
Definition einer Dummyvariable:

$$sev\_eight = \begin{cases} 1 & \text{falls } t > 1970 \text{ and } t < 1991, \\ 0 & \text{falls } t \leq 1970. \end{cases}$$

(Siehe Appendix A.1 für das R-Programm und Befehle zum Erstellen der Dummyvariablen.)

Statisches Regressionsmodell mit Dummy (Ergebnisse in Tabelle 2.4.):

$$inf_t = \beta_0 + \beta_1\, alr_t + \beta_2\, sev\_eight_t + u_t, \quad t = 1950, 1953, \ldots, 2024.$$

Interpretation von Schätz-Modell und Schätzergebnissen aus Tabelle 2.4: Vorausgesetzt, die Annahmen TS.1 bis TS.6 gelten, zeigt sich, dass sich die Phillips-Kurve in den 70iger und 80iger Jahren und noch-

mals in den 90iger und 00er Jahren gegenüber den 50iger und 60iger Jahren verschoben hat.

Tabelle 2.4.: Phillips-Kurve und Shift-Dummy

```
Call:
lm(formula = INFL_DE ~ AL_DE + sev_eight + nin_zero, data = window(data_phillips_de_zoo,
    start = 1950, end = 2024))

Residuals:
    Min      1Q  Median      3Q     Max
-3.1643 -0.9110 -0.1629  0.4518  7.1124

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.12407    0.44375   7.040 1.06e-09 ***
AL_DE       -0.25267    0.07699  -3.282 0.001609 **
sev_eight    2.13886    0.54177   3.948 0.000185 ***
nin_zero     0.62617    0.53053   1.180 0.241888
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.616 on 70 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.2827,Adjusted R-squared:  0.2519
F-statistic: 9.195 on 3 and 70 DF,  p-value: 3.305e-05

> SelectCritEViews(phillips_de_str_lm)
     aic      hq       sc
[1,] 3.850531 3.900213 3.975075
```

Daten: siehe Appendix A.1 für das R-Programm.

2.5.2. Mehrere Untergruppen binärer Merkmale

Beispiel: Lohngleichung mit Dummy-Gruppe Geschlecht – Ehestatus

Ein Arbeitnehmer ist männlich oder weiblich und verheiratet oder nicht verheiratet $\Longrightarrow$ 4 Untergruppen:

1. weiblich und nicht verheiratet,
2. weiblich und verheiratet,
3. männlich und nicht verheiratet,
4. männlich und verheiratet.

Vorgehen:

- Wähle eine Untergruppe als **Referenzgruppe**, z.B: weiblich und nicht verheiratet.
- Generiere primäre Dummy-Variablen `Female, Married`.
- Definiere Dummyvariablen für die kombinierten Untergruppen. Z.B. in R :
  - `femmarr <- female * married`,
  - `malesing <-(1-female) * (1- married)`,
  - `malemarr <-(1-female) * married`.
- Geschätztes Modell inkl. dieser Dummy-Gruppe in Tabelle 2.5.

Tabelle 2.5.: Lohnregression mit Dummy-Gruppe

```
Call:
lm(formula = log(wage) ~ 1 + femmarr + malesing + malemarr +
    educ + exper + I(exper^2) + tenure + I(tenure^2))

Residuals:
     Min       1Q   Median       3Q      Max
-1.89697 -0.24060 -0.02689  0.23144  1.09197

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.2110279  0.0966445   2.184   0.0294 *
femmarr     -0.0879174  0.0523481  -1.679   0.0937 .
malesing     0.1103502  0.0557421   1.980   0.0483 *
malemarr     0.3230259  0.0501145   6.446 2.64e-10 ***
educ         0.0789103  0.0066945  11.787  < 2e-16 ***
exper        0.0268006  0.0052428   5.112 4.50e-07 ***
I(exper^2)  -0.0005352  0.0001104  -4.847 1.66e-06 ***
tenure       0.0290875  0.0067620   4.302 2.03e-05 ***
I(tenure^2) -0.0005331  0.0002312  -2.306   0.0215 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3933 on 517 degrees of freedom
Multiple R-squared:  0.4609,Adjusted R-squared:  0.4525
F-statistic: 55.25 on 8 and 517 DF,  p-value: < 2.2e-16

> SelectCritEViews(ols_wage_1)
     aic       hq       sc
[1,] 0.9884228 1.016998 1.061403
```

Anmerkungen: Daten: `wage1.txt`. Quelle: Wooldridge (2009), siehe Appendix A.7 für das R-Programm.

Interpretation zu Tabelle 2.5:

- Verheiratete Frauen verdienen um ca. 8.8% weniger als nicht verheiratete Frauen. Der Effekt ist allerdings nur auf dem 10% Signifikanzniveau signifikant (bei einem zweiseitigen Test).

- Der Lohnunterschied zwischen verheirateten Männer und Frauen ist ca. $32.3 - (-8.8) = 41.1\%$. Ein t-Test geht hier nicht direkt. (Abhilfe: Neue Regression mit einer der beiden Untergruppen als Referenzgruppe.)

Beachte: Verwendung von Dummies für *alle* Untergruppen ist nicht möglich, da die vollständige Dummy-Gruppe kollinear ist — außer es ist keine Konstante im Regressionsmodell.

2.5.3. Ordinale Information

Beispiel: Ranglisten von Universitäten

Der Qualitätsabstand zwischen Rang 1 und 2 und Rang 11 und 12 kann sehr unterschiedlich sein. Deshalb sollte man eine Rangvariable **nicht** direkt in einer Regression verwenden. Stattdessen muss man für jede bis auf eine Universität eine Dummyvariable definieren, auch wenn dadurch mehr Parameter zu schätzen sind.

Beachte: Der Koeffizient jeder einzelnen Dummy gibt den gemessenen Unterschied im Achsenabstand zwischen der gegebenen Universität und der Referenzuniversität an.

Falls zu viele Ränge vorliegen und damit zu viele Parameter zu schätzen wären, bildet man Untergruppen, z.B. alle Unis mit Rang 20-30.

2.5.4. Interaktionsterme mit Dummyvariablen

Interaktion zwischen Dummyvariablen

Interaktion zwischen Dummyvariablen lässt sich nutzen, um direkt Untergruppen zu definieren (s.o.).

- **Beachte**, dass der Vergleich von Untergruppen nur gelingt, wenn die jeweiligen Dummywerte richtig eingesetzt werden.
- Z.B. sind im Modell die Dummies $male$ und $married$ und deren Produkt enthalten:

$$y = \beta_0 + \delta_1 male + \delta_2 married + \delta_3 male \cdot married + \dots.$$

- Vergleich zwischen (männlich, verheiratet) und (männlich, nicht verheiratet) mit:

$$\begin{aligned} &E(y|male = 1, married = 1) - E(y|male = 1, married = 0) \\ &\quad = \beta_0 + \delta_1 + \delta_2 + \delta_3 + \ldots - (\beta_0 + \delta_1 + \ldots) \\ &\quad = \delta_2 + \delta_3. \end{aligned}$$

Interaktion mit quantitativen Variablen

Interaktion mit quantitativen Variablen lässt unterschiedliche Steigungsparameter für verschiedene Gruppen zu, z.B.:

$$y = \beta_0 + \beta_1 D + \beta_2 x_1 + \beta_3 (x_1 \cdot D) + u.$$

Beachte: Nunmehr gibt β_1 den Unterschied zwischen den beiden Gruppen *nur* für $x_1 = 0$ an. Ist $x_1 \neq 0$, ist der Unterschied

$$\begin{aligned} E(y|D=1,x_1) - E(y|D=0,x_1) \\ &= \beta_0 + \beta_1 \cdot 1 + \beta_2 x_1 + \beta_3 (x_1 \cdot 1) - (\beta_0 + \beta_2 x_1) \\ &= \beta_1 + \beta_3 x_1. \end{aligned}$$

Selbst wenn β_1 negativ ist, kann der Gesamteffekt positiv sein!

Beispiel: Lohngleichung mit Interaktionstermen

Sind *returns to schooling* geschlechtsabhängig? Untersuchung anhand der Schätzergebnisse in Tabelle 2.6.

Tabelle 2.6.: Lohnregression mit Interaktionen

```
Call:
lm(formula = log(wage) ~ 1 + female + educ + exper + I(exper^2) +
    tenure + I(tenure^2) + I(female * educ))

Residuals:
     Min       1Q   Median       3Q      Max
-1.83264 -0.25261 -0.02374  0.25396  1.13584

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       0.3888060  0.1186871   3.276  0.00112 **
female           -0.2267886  0.1675394  -1.354  0.17644
educ              0.0823692  0.0084699   9.725  < 2e-16 ***
exper             0.0293366  0.0049842   5.886 7.11e-09 ***
I(exper^2)       -0.0005804  0.0001075  -5.398 1.03e-07 ***
tenure            0.0318967  0.0068640   4.647 4.28e-06 ***
I(tenure^2)      -0.0005900  0.0002352  -2.509  0.01242 *
I(female * educ) -0.0055645  0.0130618  -0.426  0.67028
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4001 on 518 degrees of freedom
Multiple R-squared:  0.441,Adjusted R-squared:  0.4334
F-statistic: 58.37 on 7 and 518 DF,  p-value: < 2.2e-16

> SelectCritEViews(ols_wage_2)
     aic     hq      sc
[1,] 1.02089 1.04629 1.085761
```

Anmerkungen: Daten:`wage1.txt`. Quelle: Wooldridge (2009), siehe Appendix A.7 für das R-Programm.

Testen auf Unterschiede zwischen den Gruppen

- Tests auf Unterschiede zwischen den Gruppen können mit Hilfe des F-Tests gemacht werden.

- **Chow-Test**: Der Chow-Test ermöglicht zu testen, ob irgendein Unterschied zwischen zwei Gruppen existiert, d.h. ob entweder der Achsenabschnitt oder/und mindestens ein Steigungsparameter verschieden ist.

  Illustration:

  $$y = \beta_0 + \beta_1 D + \beta_2 x_1 + \beta_3 (x_1 \cdot D) + \beta_4 x_2 + \beta_5 (x_2 \cdot D) + u,$$

  Hypothesen:

  $\mathsf{H}_0 : \ \beta_1 = \beta_3 = \beta_5 = 0$ vs. $\mathsf{H}_1 : \ \beta_1 \neq 0$ oder $\beta_3 \neq 0$ oder $\beta_5 \neq 0$.

Durchführung des F-Tests:

- Schätze für jede Gruppe l die Regressionsgleichung

$$y = \beta_{0l} + \beta_{2l} x_1 + \beta_{4l} x_2 + u, \qquad l = 1, 2$$

  und notiere SSR_1 und SSR_2.

- Schätze diese Gleichung für beide Gruppen zusammen und notiere SSR.

- Berechne die F-Statistik:

$$F = \frac{\text{SSR} - (\text{SSR}_1 + \text{SSR}_2)}{\text{SSR}_1 + \text{SSR}_2} \, \frac{T - 2(k+1)}{(k+1)},$$

  wobei die Zahl der Freiheitsgerade für die F-Verteilung gerade $k+1$ und $T - 2(k+1)$ sind.

Zu lesen: Chapter 7 (ohne Sections 7.5 und 7.6) in Wooldridge (2009).

3. Trends und Saisonalität

Beispiel: Bruttonationaleinkommen Vierteljahresdaten

Abbildung 3.1 (=Abbildung 1.3) zeigt die Entwicklung des deutschen Bruttonationaleinkommens und weist u.a. folgende Auffälligkeiten auf:

- Trend: langsame gleichmäßige Veränderung des durchschnittlichen Niveaus einer Zeitreihe – langfristige Entwicklung,
- Saisonalität: zyklisch sich wiederholende Muster.

Ignorieren von Trends und/oder Saisonalität führt im Allgemeinen zu verzerrten Parameterschätzungen bzw. sogar zu irreführenden Regressionsbeziehungen!

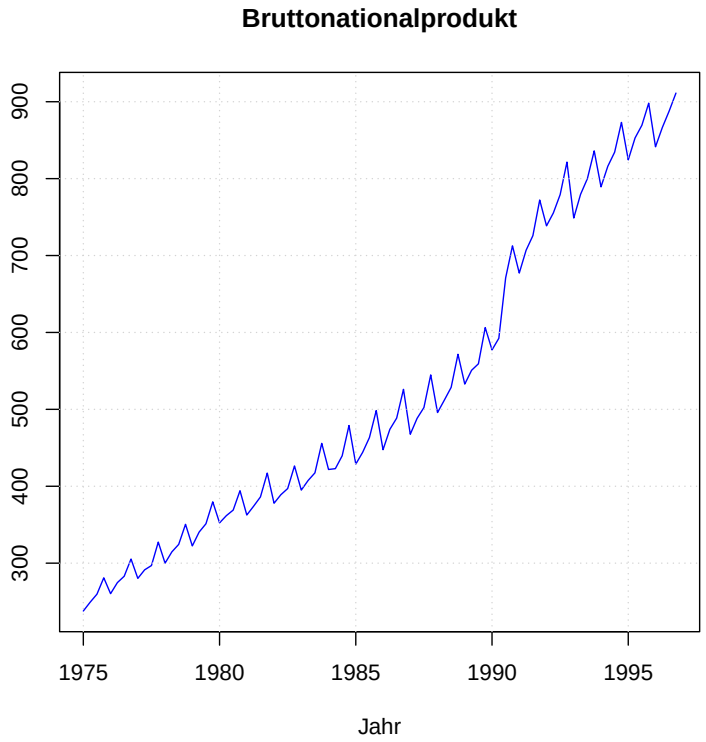

Abbildung 3.1.: Vierteljahresdaten des deutschen Bruttonationaleinkommens, Daten und R-Programm siehe Abbildung 1.3

3.1. Arten von Trends

- Deterministischer Trend:
  Einfachster Fall: linearer Zeittrend:

$$y_t = \alpha_0 + \alpha_1 t + e_t, \quad t = 1, 2, \ldots,$$

  so dass $E[y_t] = \alpha_0 + \alpha_1 t$, wenn gilt $E[e_t] = 0$.

  In der Praxis sind die Fehler häufig autokorreliert, d.h. $Cov(e_t, e_{t-k}) \neq 0$ für einige $k = \pm 1, \pm 2, \ldots$. Siehe hierzu Kapitel 4.

- Stochastischer Trend: wird in Kapitel 6 erklärt.

Später werden wir sehen, dass es für die Modellierung wichtig ist, den geeigneten Trend zu wählen.

3.2. Funktionale Formen von deterministischen Trends

- Linearer Trend,
- exponentieller Trend,
- quadratischer Trend,
- weitere funktionale Formen.

Notation

- Eine Zufallsvariable y_t ist i.i.d.(μ, σ^2) verteilt, wenn sie für verschiedene t unabhängig und identisch verteilt ist (**i**ndependently and **i**dentically **d**istributed):

$$y_t \sim \text{i.i.d.}(\mu, \sigma^2).$$

  Die i.i.d. Annahme enthält die Annahme einer Zufallsstichprobe für y_t.

- Differenzenoperator:

$$\Delta y_t \equiv y_t - y_{t-1}.$$

3.2.1. Linearer Trend

Wie bereits gesehen, lässt sich ein linearer Trend wie folgt abbilden:

$$y_t = \alpha_0 + \alpha_1 t + e_t, \quad E[e_t] = 0, \quad t = 1, 2, \ldots.$$

Interpretation

$$E[y_t] = \alpha_0 + \alpha_1 t$$

und somit:

$$\Delta E[y_t] = E[y_t] - E[y_{t-1}] = \alpha_1.$$

Alternativ (für $\Delta e_t = 0$):

$$\Delta y_t = \alpha_1.$$

Der Parameter α_1 gibt die Veränderung des unbedingten Erwartungswerts pro Zeiteinheit an, z.B. den absoluten Zuwachs des realen BIP,

siehe Appendix A.3 für das R-Programm, das linearen Trend und entsprechende Residuen schätzt und graphisch darstellt.

Schätzung

Ist zusätzlich für die e_t die i.i.d. Annahme erfüllt, erfüllt das lineare Trendmodell die Annahmen TS.1 bis TS.5.

Für den Fall autokorrelierter Fehler siehe Kapitel 8.

Es ist möglich, zusätzlich zu einem Zeittrend t weitere Regressoren aufzunehmen, siehe weiter unten.

3.2.2. Exponentieller Zeittrend

Die folgende Spezifikation bildet einen exponentiellen Trend ab:

$$\log(y_t) = \beta_0 + \beta_1 t + e_t, \quad e_t \sim \text{i.i.d.}(0, \sigma^2), \quad t = 1, 2, \ldots.$$

Interpretation

$$y_t = e^{\beta_0 + \beta_1 t} e^{e_t}.$$

Und somit erhält man (siehe exakte Effekte in log-level Modellen in Abschnitt 2.4.2):

$$\begin{aligned} E[y_t] &= e^{\beta_0 + \beta_1 t} E[e^{e_t}], \\ \frac{E[y_t] - E[y_{t-1}]}{E[y_{t-1}]} &= e^{\beta_1} - 1, \end{aligned}$$

da $E[e^{e_t}] = E[e^{e_{t-1}}]$ aufgrund der i.i.d. Annahme.

Somit entspricht $100\beta_1$ approximativ der **Wachstumsrate** von y_t pro Periode in Prozent.

Schätzung

Wie bei statischer Regression. Siehe Appendix A.3 für das R-Programm, das für das deutsche reale BIP einen exponentiellen Trend und entsprechende Residuen schätzt und graphisch darstellt.

3.2.3. Quadratischer Trend

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + e_t, \quad e_t \sim \text{i.i.d.}(0, \sigma^2), \quad t = 1, 2, \ldots.$$

Interpretation

$$E[y_t] = \alpha_0 + \alpha_1 t + \alpha_2 t^2.$$

Nimmt man für den Augenblick an, dass der Zeitindex t stetig ist, ergibt sich

$$\frac{dE[y_t]}{dt} = \alpha_1 + 2\alpha_2 t.$$

Der Zeittrend ist ansteigend, falls $\alpha_1, \alpha_2 > 0$. Und sonst?

Schätzung

Wie bei statischer Regression. Siehe Appendix A.3 für das R-Programm, das für das deutsche reale BIP einen quadratischen Trend und entsprechende Residuen schätzt und graphisch darstellt.

3.3. Trendbehaftete Variablen in der Regressionsanalyse

Schätzeigenschaften und Interpretation hängen vom wahren datengenerierenden Prozess (Grundgesamtheit) ab. Der allgemeine Fall ist (ausgehend von linearen Trends):

$$y_t = \alpha_0 + \alpha t + \tilde{y}_t, \tag{3.1}$$

$$x_t = \delta_0 + \delta t + \tilde{x}_t, \tag{3.2}$$

wobei die $\tilde{y}_t$ und $\tilde{x}_t$ die Abweichungen vom jeweiligen linearen Trend darstellen.

3.3.1. Fall I

Es gibt einen linearen Zusammenhang zwischen den trendbehafteten Variablen:

$$y_t = \beta_0 + \beta_1 x_t + u_t, \quad u_t \sim \text{i.i.d.}(0, \sigma^2).$$

Schätzung

Dann funktioniert alles wie gewohnt, solange außerdem Annahmen TS.2 (Keine Kollinearität, Abschnitt 2.2.1) und TS.3 (Strikte Exogenität, Abschnitt 2.2.1) erfüllt sind.

Darstellung in Abweichungen vom Trend

Der Zusammenhang lässt sich auch in den Abweichungen vom Trend darstellen, indem man (3.2) und (3.1) in die Regressionsgleichung einsetzt und nach $\tilde{y}_t$ auflöst:

$$\tilde{y}_t = (\beta_0 + \beta_1\delta_0 - \alpha_0) + \beta_1\tilde{x}_t + (\beta_1\delta - \alpha)\, t + u_t.$$

Man sieht, dass man bei der Schätzung in den Abweichungen vom Trend den Zeittrend mit in die Regression aufnehmen muss, um Verzerrungen zu vermeiden, **außer** es gilt:

$$\beta_1\delta - \alpha = 0.$$

Kleiner Ausblick

Der Trendterm $(\beta_1\delta - \alpha)\, t$ lässt sich auch schreiben als:

$$\begin{pmatrix} \beta_1 & -1 \end{pmatrix} \begin{pmatrix} \delta t \\ \alpha t \end{pmatrix}.$$

Wenn also das Produkt der beiden Vektoren $\begin{pmatrix} \beta_1 & -1 \end{pmatrix}$ und $\begin{pmatrix} \delta t & \alpha t \end{pmatrix}'$ gerade 0 ergibt, dann eliminiert der Vektor $\begin{pmatrix} \beta_1 & -1 \end{pmatrix}$ den Trend aus der Beziehung der Abweichungen.

Später, in Kapitel 10, werden wir den Vektor $\begin{pmatrix} \beta_1 & -1 \end{pmatrix}$ als *Kointegrationsvektor* bezeichnen, wenn er zumindest einen Teil des Trends (genauer: den stochastischen Trend) aus der Beziehung der Abweichungen eliminieren kann.

3.3.2. Fall II

Es gibt keinen linearen Zusammenhang zwischen den trendbehafteten Variablen (d.h. $\beta_1 = 0$), aber zwischen den Abweichungen:

$$\tilde{y}_t = \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{x}_t + u_t, \quad u_t \sim \text{i.i.d.}(0, \sigma^2). \tag{3.3}$$

Auflösen von (3.2) und (3.1) nach den Abweichungen vom Trend und Einsetzen in (3.3) ergibt:

$$\begin{aligned} y_t &= \alpha_0 + \alpha t + \tilde{\beta}_0 + \tilde{\beta}_1 \left(x_t - \delta_0 - \delta t\right) + u_t \\ &= \tilde{\beta}_0 + \alpha_0 - \tilde{\beta}_1 \delta_0 + \tilde{\beta}_1 x_t + \left(\alpha - \tilde{\beta}_1 \delta\right) t + u_t. \end{aligned}$$

Aus **Einführung in die Ökonometrie**, Abschnitt 3.3 ist bekannt, dass bei Vernachlässigung einer Variable die Parameterschätzung der anderen Variablen verzerrt sein kann.

In Matrixschreibweise lässt sich das Modell schreiben als:

$$\mathbf{y} = \mathbf{X}_x \begin{pmatrix} \tilde{\beta}_0 + \alpha_0 - \tilde{\beta}_1 \delta_0 \\ \tilde{\beta}_1 \end{pmatrix} + \mathbf{x}\left(\alpha - \tilde{\beta}_1 \delta\right) + \mathbf{u}, \quad \mathbf{X}_x = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_t \\ \vdots & \vdots \\ 1 & x_T \end{pmatrix}, \mathbf{x} = \begin{pmatrix} 1 \\ \vdots \\ t \\ \vdots \\ T \end{pmatrix}.$$

Entsprechend erhält man bei Vernachlässigung von Regressor $\mathbf{x}$:

$$\widehat{\tilde{\boldsymbol{\beta}}} = \tilde{\boldsymbol{\beta}} + \left(\mathbf{X}_x'\mathbf{X}_x\right)^{-1}\mathbf{X}_x'\mathbf{x}\left(\alpha - \tilde{\beta}_1\delta\right) + \left(\mathbf{X}_x'\mathbf{X}_x\right)^{-1}\mathbf{X}_x'\mathbf{u}. \quad (3.4)$$

Man beachte, dass wegen $t = \frac{1}{\delta}x_t - \frac{\delta_0}{\delta} - \frac{1}{\delta}\tilde{x}_t$ nach (3.2) (wobei $\delta \neq 0$ vorausgesetzt wurde) gilt:

$$\left(\mathbf{X}_x'\mathbf{X}_x\right)^{-1}\mathbf{X}_x'\mathbf{x} = \begin{pmatrix} \widehat{\begin{pmatrix} -\delta_0 \\ \delta \end{pmatrix}} \\ \widehat{\begin{pmatrix} 1 \\ \delta \end{pmatrix}} \end{pmatrix}.$$

Es können verschiedene Fälle auftreten. Für β_1 ergibt sich:

- Es gilt $\alpha - \tilde{\beta}_1 \delta = 0$, so dass der Zeittrend aus der Regression herausfällt und es zu keiner Verzerrung kommt.

- Es gilt $\alpha - \tilde{\beta}_1 \delta \neq 0$ und $\alpha \neq 0$, $\delta \neq 0$. Dann beträgt die Verzerrung für $\widehat{\tilde{\beta}_1}$ für eine gegebene Stichprobe:

$$\widehat{\left(\frac{1}{\delta}\right)} \left(\alpha - \tilde{\beta}_1 \delta\right).$$

  Im Fall, dass $1/\delta$ gerade fehlerfrei geschätzt werden würde, ergäbe sich eine Verzerrung von:

$$\frac{\alpha}{\delta} - \tilde{\beta}_1.$$

- Im Unterschied zu vorherigem Fall habe y_t **keinen linearen Trend**, d.h. $\alpha = 0$. Dann könnte die Verzerrung gerade so groß wie der zu

schätzende Parameter selbst sein, nämlich $\tilde{\beta}_1$!! (Intution?)

- Weist hingegen der Regressor x_t **keinen linearen Trend** auf, d.h. $\delta = 0$, dann beträgt die Verzerrung gerade:

$$\left(\mathbf{X}_x'\mathbf{X}_x\right)^{-1}\mathbf{X}_x'\mathbf{x}\alpha.$$

  Kurz: **Die Vernachlässigung des linearen Zeittrends in einer statischen Regression zwischen den trendbehafteten Variablen kann zu Verzerrungen führen**.

Mit dem R-Programm in Appendix A.8 kann der Effekt einer fehlspezifizierten Regression anhand einer Simulation illustriert werden.

3.3.3. Fall III

Es gibt weder einen linearen Zusammenhang zwischen den trendbehafteten Variablen, noch zwischen deren Abweichungen, d.h. $\beta_1 = 0$ und $\tilde{\beta}_1 = 0$.

Einsetzen von $\tilde{\beta}_1 = 0$ in (3.4) ergibt:

$$\widehat{\tilde{\boldsymbol{\beta}}} = \begin{pmatrix} \tilde{\beta}_0 \\ 0 \end{pmatrix} + \left(\mathbf{X}_x'\mathbf{X}_x\right)^{-1}\mathbf{X}_x'\mathbf{x}\alpha + \left(\mathbf{X}_x'\mathbf{X}_x\right)^{-1}\mathbf{X}_x'\mathbf{u},$$

so dass sich im Allgemeinen $\widehat{\tilde{\beta}}_1 \neq 0$ ergibt!! In diesem Fall liefert eine Regression von y_t auf x_t (ohne Zeittrend) einen Zusammenhang, *selbst wenn es keinen* ökonomischen oder anderen Zusammenhang gibt. In diesem Fall spricht man von einer **Scheinregression** (spurious regression).

Mit dem R-Programm in Appendix A.8 kann der Effekt einer Scheinregression anhand einer Simulation illustriert werden.

Zusammenfassung

- **Die Vernachlässigung des linearen Zeittrends in einer statischen Regression zwischen den trendbehafteten Variablen kann zu Verzerrungen bzw. im schlimmsten Fall zu einer Scheinregression führen**.
- Man nimmt deshalb im Fall mindestens einer trendbehafteten Variable am besten immer einen Trendterm als Regressor auf und schätzt im allgemeinen Fall bei linearem Trend das Modell:

$$y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + d\,t + u_t.$$

Wann kann man den Trendterm $d\,t$ weglassen? Wie lautet jeweils die Interpretation der Parameter β_1 bis β_k?

- Wie lautet die Interpretation der Parameter β_1 bis β_k, falls das Modell in den Abweichungen von den Trends geschätzt wird:

$$\tilde{y}_t = \beta_0 + \beta_1 \tilde{x}_{t1} + \ldots + \tilde{\beta}_k x_{tk} + d\,t + u_t?$$

Interpretation des Bestimmtheitsmaßes

Die Interpretation des Bestimmtheitsmaßes R^2 ist bei Regressionen mit trendbehafteten Variablen nicht sinnvoll: Bei Berücksichtigung einer Konstante ist:

$$R^2 = 1 - \frac{\mathrm{SSR}}{\mathrm{SST}} = 1 - \frac{\mathrm{SSR}/T}{\mathrm{SST}/T}.$$

Während bei homoskedastischer Varianz der Varianzschätzer SSR$/T$ sich mit zunehmenden Stichproben nicht systematisch verändern sollte, ist dies bei SST$/T$ aufgrund des Trends der Fall. Damit tendiert R^2 mit zunehmender Stichprobengröße gegen 1. Der Schätzer SST$/T$ überschätzt $Var(y_t)$, da SST die Trendkomponente $d\,t$ enthält, jedoch $Var(y_t) = E\left[(y_t - E[y_t])^2\right]$ nicht!

Abhilfe: Partialling out: Man regressiert beide Variablen zunächst auf einen Zeittrend und regressiert dann die jeweiligen Abweichungen aufeinander und berechnet das Bestimmtheitsmaß. Zur Berechnung des bereinigten Bestimmtheitsmaßes siehe Section 10.5 in Wooldridge (2009).

3.4. Saisonalität

- **Saisonmuster**
  - jahreszeitliche Schwankungen in Monats-, Quartalsdaten,
  - wochentagsabhängige Schwankungen in Wochendaten,
  - tageszeitabhängige Schwankungen bei Stundendaten.
- Saisonmuster können durch **Saisonbereinigungsverfahren** mehr oder weniger herausgefiltert werden. Es gibt eine Reihe von Saisonbereinigungsverfahren, die je nach Situation unterschiedlich gut funktionieren, siehe z.B. Neusser (2009).
- Im Prinzip ist es am besten, die Saisonbereinigung selbst durch-

zuführen. Dies ist jedoch nur möglich, wenn **nicht-saisonbereinigte Daten** zur Verfügung gestellt werden. In den USA ist dies selten der Fall.

Einfaches Saisonbereinigungsverfahren

Ein sehr einfaches Saisonbereinigungsverfahren besteht darin, für jede (bis auf eine) Saison eine Dummyvariable zu definieren (vgl. Abschnitt 2.5), sogenannte **Saisondummies**.

Sieh Appendix A.3 für das R-Programm, das linearen Trend und Saisondummies schätzt und graphisch darstellt.

Beispiel für jahreszeitliche Schwankungen in Quartalsdaten (Referenzquartal: Frühjahr):

$$S_t = \begin{cases} 1 & \text{falls } t \text{ ein Sommerquartal bezeichnet,} \\ 0 & \text{sonst,} \end{cases}$$

$$H_t = \begin{cases} 1 & \text{falls } t \text{ ein Herbstquartal bezeichnet,} \\ 0 & \text{sonst,} \end{cases}$$

$$W_t = \begin{cases} 1 & \text{falls } t \text{ ein Winterquartal bezeichnet,} \\ 0 & \text{sonst.} \end{cases}$$

Eine Regression mit Saisondummies lautet z.B.:

$$y_t = \beta_0 + \beta_1 x_t + \delta_1 S_t + \delta_2 H_t + \delta_3 W_t + u_t.$$

Interpretation: Der Parameter δ_1 gibt an, um wieviel sich die Konstante im Sommer von der Konstante im Frühjahr (dem Basisquartal) unterscheidet, etc.

Zu lesen: Section 10.5 in Wooldridge (2009).

4. Regressionsmodelle für Zeitreihendaten II: Autoregressionsmodelle - verzögerte endogene Variablen als Regressoren

4.1. Grundlagen zu stochastischen Prozessen

Bisher nur informelle Überlegungen zu stochastischen Prozessen (und Zeitreihendaten) in Abschnitt 1.4.

4.1.1. Momente

- **Erwartungswert** $E[y_t]$ von y_t
  - Kann vom Zeitpunkt t abhängen. Beispiel:

$$y_t = a_0 + a_1 t + e_t, \quad e_t \sim \text{i.i.d.}(0, \sigma^2),$$
$$E[y_t] = a_0 + a_1 t.$$

  - Kann über die Zeit konstant sein.

- **Autokovarianzfunktion**
  Für alle $k = \ldots, -2, -1, 0, 1, 2, \ldots$ gilt:

$$Cov(y_t, y_{t-k}) = E\left[(y_t - E[y_t])(y_{t-k} - E[y_{t-k}])\right].$$

Der Zusatz "Auto" wird verwendet, um zu verdeutlichen, dass sich

die Kovarianz auf zwei Zufallsvariablen in *einem* Zeitreihenprozess bezieht.

- **Autokorrelationsfunktion**

  Für alle $k = \ldots, -2, -1, 0, 1, 2, \ldots$ gilt:

$$Corr\left(y_t, y_{t-k}\right) = \frac{Cov\left(y_t, y_{t-k}\right)}{\sqrt{Var\left(y_t\right) Var\left(y_{t-k}\right)}}.$$

  Es gilt

$$-1 \leq Corr\left(y_t, y_{t-k}\right) \leq 1,$$

$$Cov\left(y_t, y_{t-k}\right) = Corr\left(y_t, y_{t-k}\right) \sqrt{Var\left(y_t\right)} \sqrt{Var\left(y_{t-k}\right)}.$$

4.1.2. Stationarität

- **Schwache Stationarität**
  Sind die Erwartungswerte und die Autokovarianzfunktion von t unabhängig, d.h.

  – $E[y_t] = \mu$,

  – $Cov(y_t, y_{t-k}) = \gamma_k$, für alle $k = \ldots, -1, 0, 1, \ldots$,

  wird ein stochastischer Prozess als **(schwach) stationär** oder **kovarianzstationär** bezeichnet.
  Im Falle eines schwach stationären Prozesses gilt für die Autokorrelationsfunktion:

$$\rho_k = \frac{\gamma_k}{\gamma_0}.$$

- **Strenge Stationariät**

  Ein stochastischer Prozess wird als **streng stationär** bezeichnet, wenn die gemeinsame Wahrscheinlichkeitsverteilung einer endlichen Menge von Zufallsvariablen $(y_{t_1}, y_{t_2}, \ldots, y_{t_m})$ identisch ist mit der gemeinsamen Verteilungsfunktion der um s Perioden verschobenen Menge von Zufallsvariablen $(y_{t_1+s}, y_{t_2+s}, \ldots, y_{t_m+s})$.

  Beachte:

  – Wooldridge (2009, Ch. 11) versteht unter *Stationarität* immer *strenge Stationarität (strict stationarity)*.

  – Im Gegensatz hierzu bezeichnet in großen Teilen der Literatur *Stationarität* lediglich *schwache Stationarität*.

  – Ist ein stochastischer Prozess streng stationär und existieren des-

sen Erwartungswert und Varianz, so ist er immer auch schwach stationär. Die Umkehrung gilt jedoch nicht. Wieso?

- Beispiel eines streng stationären Prozesses mit endlichem Erwartungswert und Varianz: $y_t \sim i.i.d.(0, \sigma^2)$.

- Um Verwirrung zu vermeiden, wird in diesen Folien die genaue Art der Stationarität angegeben, falls dies nötig ist.

- **Weak Dependence**
  Ein stochastischer Prozess wird als "weakly dependent" bezeichnet, wenn die stochastischen Abhängigkeiten zwischen den Zufallsvariablen y_{t+k} und y_t mit zunehmendem k "schnell genug" abnehmen und für $k \to \infty$ Null werden. "Weak Dependence" bedeutet also, dass für große k die Zufallsvariablen y_{t+k} und y_t "beinahe" unabhängig sind.

Stochastische Abhängigkeit liegt vor, wenn die Autokovarianz ungleich Null ist. (Die Umkehrung gilt nicht. Warum?) Voraussetzung für "Weak Dependence" ist also, dass die Autokovarianzen mit zunehmendem k schnell genug gegen Null gehen.

Das Konzept der "Weak Dependence" wird in Abschnitt 5.3 wichtig, wenn es um die Bestimmung der (asymptotischen) Schätzeigenschaften bei autoregressiven Prozessen geht.

Ein stochastischer Prozess kann weakly dependent sein ohne jedoch schwach stationär zu sein und umgekehrt.

- **Trend-Stationariät**
  Ist der stochastische Prozess der Abweichungen um einen deterministischen Zeittrend stationär und weakly dependent, wird der trendbehaftete stochastische Prozess als **trend-stationär** bezeichnet.

4.1.3. Störterme

- **Weißes Rauschen** (White noise)
  $e_t \sim WN(0, \sigma^2)$ bedeutet:
  - $E[e_t] = 0, \quad t = \ldots, -2, -1, 0, 1, 2, \ldots$, der Mittelwert von e_t ist für jede Periode 0,
  - $Var(e_t) = E[e_t^2] = \sigma^2, \quad t = \ldots, -2, -1, 0, 1, 2, \ldots$, die Varianz von e_t, d.h. die erwartete quadratische Abweichung, ist konstant über die Zeit (es liegt keine Heteroskedastie vor),
  - $Cov(e_t, e_{t-k}) = 0$ für $k \neq 0$, die Kovarianz zwischen verschiedenen Perioden ist 0, d.h. es hilft nichts, e_{t-k} zu beobachten, um mehr über die Wahrscheinlichkeit zu wissen, dass eine Realisation von e_t in einem bestimmten Intervall liegt.

Weißes Rauschen ist **schwach stationär**.

- **Gaussches Weißes Rauschen** (Gaussian White noise)
  $e_t \sim$ i.i.d.$N(0, \sigma^2)$ bedeutet:
  - $E[e_t] = 0, \quad t = \ldots, -2, -1, 0, 1, 2, \ldots,$
  - $Var(e_t) = E[e_t^2] = \sigma^2, \quad t = \ldots, -2, -1, 0, 1, 2, \ldots,$
  - $Cov(e_t, e_{t-k}) = 0$ für $k \neq 0,$
  - $e_t \sim N(0, \sigma^2)$, d.h. damit sind die e_t und e_s, $t \neq s$, unabhängig und identisch verteilt, kurz es gilt dann auch $e_t \sim$ i.i.d.$(0, \sigma^2)$.

Gaussches Weißes Rauschen ist deshalb **streng stationär**.

4.2. Autoregressive Prozesse erster Ordnung

4.2.1. Definitionen

- **Lineare Differenzengleichung erster Ordnung**
  Zur Erinnerung: (Deterministische, homogene) lineare Differenzengleichungen (DGL) erster Ordnung (mit konstantem Koeffizienten, siehe Modul **Methoden der VWL**):

$$x_t = a x_{t-1}.$$

  Lösung:

$$x_t = a^t x_0.$$

- **Autoregressiver Prozess erster Ordnung**
  Ein stochastischer Prozess $\{y_t : t = \ldots, -2, -1, 0, 1, 2, \ldots\}$ heißt **autoregressiver Prozess erster Ordnung** (AR(1)-Prozess), wenn er folgende *stochastische lineare Differenzengleichung erster Ordnung* erfüllt:
  (siehe Modul **Methoden der VWL**):

$$y_t = \alpha y_{t-1} + e_t, \tag{4.1}$$

  und die e_t's weißes Rauschen sind, also $e_t \sim WN(0, \sigma^2)$.

  Anstelle von $e_t \sim WN(0, \sigma^2)$ kann auch die strengere Annahme $e_t \sim i.i.d.(0, \sigma^2)$ vorausgesetzt werden, so wie in Wooldridge (2009, Equation (11.1)) und später in Abschnitt 5.3.1.

4.2.2. "Lösung" eines AR(1) Prozesses

Einsetzen der einfach verzögerten Gleichung

$$y_{t-1} = \alpha y_{t-2} + e_{t-1}$$

und der mehrfach verzögerten Gleichung in die autoregressive Gleichung für y_t ergibt:

$$\begin{aligned}
y_t &= \alpha\left(\alpha y_{t-2} + e_{t-1}\right) + e_t = \alpha^2 y_{t-2} + \alpha e_{t-1} + e_t \\
&= \alpha^2\left(\alpha y_{t-3} + e_{t-2}\right) + \alpha e_{t-1} + e_t \\
&= \alpha^3 y_{t-3} + \alpha^2 e_{t-2} + \alpha e_{t-1} + e_t \\
&\;\;\vdots \\
&= \alpha^t y_0 + \alpha^{t-1} e_1 + \alpha^{t-2} e_2 + \cdots + \alpha e_{t-1} + e_t \\
&= \alpha^t y_0 + \sum_{j=0}^{t-1} \alpha^j e_{t-j}. \qquad (4.2)
\end{aligned}$$

- Selbst wenn y_0 eine nicht-stochastische Größe ist, ist die Variable y_t für jedes $t > 0$ eine Zufallsvariable.

- Ohne die gewichtete Summe des Weißen Rauschens erhält man gerade die Lösung einer deterministischen linearen Differenzengleichung erster Ordnung.

- Die **Stabilitätseigenschaften** eines AR(1)-Prozesses entsprechen deshalb denen einer linearen DGL 1. Ordnung.

- Im Fall $\alpha = 1$ wird der AR(1)-Prozess als **Random Walk** bezeichnet (vgl. Abschnitt 4.2.6).

4.2.3. Stochastische Eigenschaften von AR(1)-Prozessen

- Es wird angenommen, dass y_0 und e_t, $t = 1, 2, \ldots$ stochastisch unabhängig sind.

- **Erwartungswert**

$$E[y_t] = E\left[\alpha^t y_0 + \sum_{j=0}^{t-1} \alpha^j e_{t-j}\right]$$
$$= \alpha^t E\left[y_0\right] + \sum_{j=0}^{t-1} \alpha^j E\left[e_{t-j}\right] = \alpha^t E\left[y_0\right].$$

  Der Erwartungswert ist 0 für alle t, falls der Startwert $y_0 = 0$ ist bzw. dessen Erwartungswert $E[y_0] = 0$ ist.

- **Varianz - allgemeiner Fall**

$$Var(y_t) = \alpha^{2t} Var(y_0) + Var\left(\sum_{j=0}^{t-1} \alpha^j e_{t-j}\right) + 2\,Cov\left(\alpha^t y_0, \sum_{j=0}^{t-1} \alpha^j e_{t-j}\right)$$

$$Var(y_t) \overset{e_t \sim \text{WN}, y_0 \text{ und } e_t \text{ stoch. unabh.}}{=} \alpha^{2t} Var(y_0) + \sum_{j=0}^{t-1} Var\left(\alpha^j e_{t-j}\right)$$

$$= \alpha^{2t} Var(y_0) + \sum_{j=0}^{t-1} \alpha^{2j} Var\left(e_{t-j}\right)$$

$$\overset{e_t \sim \text{WN}}{=} \alpha^{2t} Var(y_0) + \sigma^2 \sum_{j=0}^{t-1} \alpha^{2j}.$$

  - Die Varianz hängt im allgemeinen Fall von t ab.
  - Ein AR(1)-Prozess ist ohne weitere Annahmen im allgemeinen

nicht kovarianzstationär.

- **Autokovarianzfunktion - allgemeiner Fall**

$$\begin{aligned} Cov(y_t, y_{t-1}) &= Cov(\alpha y_{t-1} + e_t, y_{t-1}) \\ &= \alpha Var(y_{t-1}) + Cov(e_t, y_{t-1}) \\ &= \alpha Var(y_{t-1}). \end{aligned}$$

Für Abstände $k \geq 0$ zwischen den Perioden erhält man:

$$\begin{aligned} Cov(y_t, y_{t-k}) &= \alpha Cov(y_{t-1}, y_{t-k}) \\ &= \alpha^k Var(y_{t-k}). \end{aligned}$$

Es gilt entsprechend:

$$Cov(y_{t+k}, y_t) = \alpha^k Var(y_t).$$

Hinweis zur "weak dependence":
AR(1) ist für $|\alpha| < 1$ weakly dependent, da die Autokovarianzen dann exponentiell schnell gegen Null gehen.

4.2.4. Stochastische Eigenschaften stationärer AR(1)-Prozesse

Man kann (4.2) umformen zu $y_t = \alpha^s y_{t-s} + \sum_{j=0}^{s-1} \alpha^j e_{t-j}$. Unter der Annahme $|\alpha| < 1$ ergibt sich für $s \to \infty$ (der Prozess läuft schon unendlich lange)

$$y_t = \sum_{j=0}^{\infty} \alpha^j e_{t-j}. \tag{4.3}$$

Da $e_t \sim WN(0, \sigma^2)$ vorausgesetzt wurde, gilt:

- **Erwartungswert**

$$E[y_t] = 0.$$

- **Varianz**

$$Var(y_t) = \sigma^2 \sum_{j=0}^{\infty} \alpha^{2j} = \frac{\sigma^2}{1-\alpha^2}.$$

- **Autokovarianzfunktion**

  Ist die Varianz über die Zeit konstant, ist damit auch die Kovarianz unabhängig von t und mit $\gamma_0 \equiv Var(y_t)$ gilt:

$$\gamma_k \equiv Cov(y_t, y_{t-k}) = \alpha^k \gamma_0.$$

- **Autokorrelationsfunktion**

$$\rho_k \equiv Corr\left(y_t, y_{t-k}\right) = \frac{Cov\left(y_t, y_{t-k}\right)}{\sqrt{Var\left(y_t\right) Var\left(y_{t-k}\right)}} = \frac{\alpha^k \gamma_0}{\sqrt{\gamma_0}\sqrt{\gamma_0}} = \alpha^k.$$

Bemerkungen:

- Der AR(1)-Prozess hat **immer von 0 verschiedene Autokorrelationen**.
- Die Autokorrelationen gehen exponentiell schnell gegen Null, da $\rho_k = \alpha^k$.
- Damit ist ein schon unendlich lange laufender AR(1)-Prozess (4.1) mit $e_t \sim WN(0, \sigma^2)$ und $|\alpha| < 1$ **schwach stationär**.
- Gilt darüber hinaus, dass der Fehlerprozess streng stationär ist, $e_t \sim i.i.d.(0, \sigma^2)$, dann ist der AR(1)-Prozess ebenfalls **streng stationär**.

4.2.5. Darstellung als unendlicher Moving-Average-Prozess

Ist der AR(1)-Prozess stationär, dann existiert die Darstellung (4.3).

Definiert man $\phi_j \equiv \alpha^j$, dann lässt sich die Darstellung (4.3) auch schreiben als:

$$y_t = \sum_{j=0}^{\infty} \alpha^j e_{t-j} = \sum_{j=0}^{\infty} \phi_j e_{t-j}. \tag{4.4}$$

Der stochastische Prozess wird als unendlicher **Moving-Average-Prozess**, kurz als **MA(∞)-Prozess** bezeichnet.

4.2.6. Random Walk

Spezieller AR(1)-Prozess mit $\alpha = 1$:

$$y_t = y_{t-1} + e_t, \quad e_t \sim WN(0, \sigma^2), \quad t = 1, 2, \ldots$$

Einsetzen von $\alpha = 1$ in die Formeln aus 4.2.3 ergibt $y_t = y_0 + \sum_{s=1}^{t} e_s$:

- Erwartungswert: $E[y_t] = E[y_0]$,
- Varianz: $Var(y_t) = \sigma^2 t + Var(y_0)$,
- Autokovarianz: $Cov(y_t, y_{t-k}) = \sigma^2(t-k)$ für $t > k \geq 0$, falls $Var(y_0) = 0$.

Also: nichtstationärer stochastischer Prozess, auch nicht "weakly dependent".

4.2.7. Bedingte Momente

- **Bedingter Erwartungswert**

$$E[y_t|y_{t-1}] = \alpha y_{t-1} + E[e_t|y_{t-1}] = \alpha y_{t-1},$$

wenn $E[e_t|y_{t-1}] = 0$ gilt. Letzteres gilt, wenn $e_t \sim i.i.d.(0, \sigma^2)$ und e_t, y_0 stochastisch unabhängig sind. Der bedingte Erwartungswert αy_{t-1} ist im Allgemeinen ungleich $E[y_t]$, unabhängig vom Parameter α.

Definiert man im einfachen Regressionsmodell den Regressor $x_t \equiv y_{t-1}$, so erhält man die bekannte Version $E[y_t|x_t] = \alpha x_t$.

- **Bedingte Varianz**

$$Var(y_t|y_{t-1}) = \sigma^2,$$

ist konstant, falls e_t weißes Rauschen ist, selbst wenn $Var(y_t)$ nicht

unabhängig von t ist.

Es gibt auch autoregressive Prozesse mit (bedingt) heteroskeastischer Fehlervarianz.

4.2.8. Deterministische Terme in AR(1)-Prozessen

AR(1)-Prozesse können auch einen Mittelwert ungleich Null haben.

- **Stationärer Fall**
  Im Fall eines stationären AR(1)-Prozesses gilt $E[y_t] = \mu$. Da $\tilde{y}_t = y_t - \mu$ Erwartungswert Null hat, ergibt sich:

$$\begin{aligned} \tilde{y}_t &= \alpha \tilde{y}_{t-1} + e_t, \\ y_t &= \mu + \alpha y_{t-1} - \alpha\mu + e_t \\ &= \underbrace{\nu}_{\mu(1-\alpha)} + \alpha y_{t-1} + e_t. \end{aligned}$$

  Sofern $|\alpha| < 1$, berechnet sich die Konstante ν des Regressionsmodells aus Mittelwert und AR-Parameter.

- **Trend-stationärer Fall**

  Ist der unbedingte Erwartungswert linear von der Zeit abhängig,

  $$E[y_t] = \mu_t = \mu_0 + \mu_1 t,$$

  erhält man ein AR(1)-Prozess um einen linearen Zeittrend.

  Ist $|\alpha| < 1$, erhält man mittels $\tilde{y}_t = y_t - \mu_t = y_t - \mu_0 - \mu_1 t$:

  $$y_t = \underbrace{\nu_0}_{\mu_0(1-\alpha)+\alpha\mu_1} + \underbrace{\nu_1}_{\mu_1(1-\alpha)} t + \alpha y_{t-1} + e_t.$$

  Falls $|\alpha| < 1$, wird der Prozess $\{y_t\}$ als trend-stationär bezeichnet.

- **Random Walk**

  Im Fall eines Random Walks erhält die Konstante eine andere Interpretation, siehe dazu Abschnitt 6.3.

4.3. Monte-Carlo-Simulationen

Ist die vollständige Spezifikation eines stochastischen Prozesses bekannt, dann lassen sich mit einem Computer beliebig viele Realisationen von diesem stochastischen Prozess erzeugen. Man führt dann sogenannte **Monte-Carlo-Simulationen** durch.

- **Vollständige Spezifikation** für AR(1)-Prozess ohne Konstante:

  1. Wahrscheinlichkeitsverteilung der Fehler $\{e_t, t = 1, \ldots, T$ (man kann nur endlich viele Beobachtungen generieren),

  2. Autoregressionsparameter α,

  3. Startwert/Vorstichprobenwert y_0 bzw. dessen Wahrscheinlichkeitsverteilung.

4.3.1. Monte-Carlo-Simulation einer Realisation

Für die folgende vollständige Spezifikation eines AR(1)-Prozesses

1. $e_t \sim$ i.i.d.$N(0,1)$,

2. $\alpha = 0.9$,

3. $y_0 = 0$.

zeigt Abbildung 4.1 eine Realisierung eines simulierten AR(1)-Prozesses mit $T = 50$ Beobachtungen. Das R-Programm findet sich in Appendix A.9.

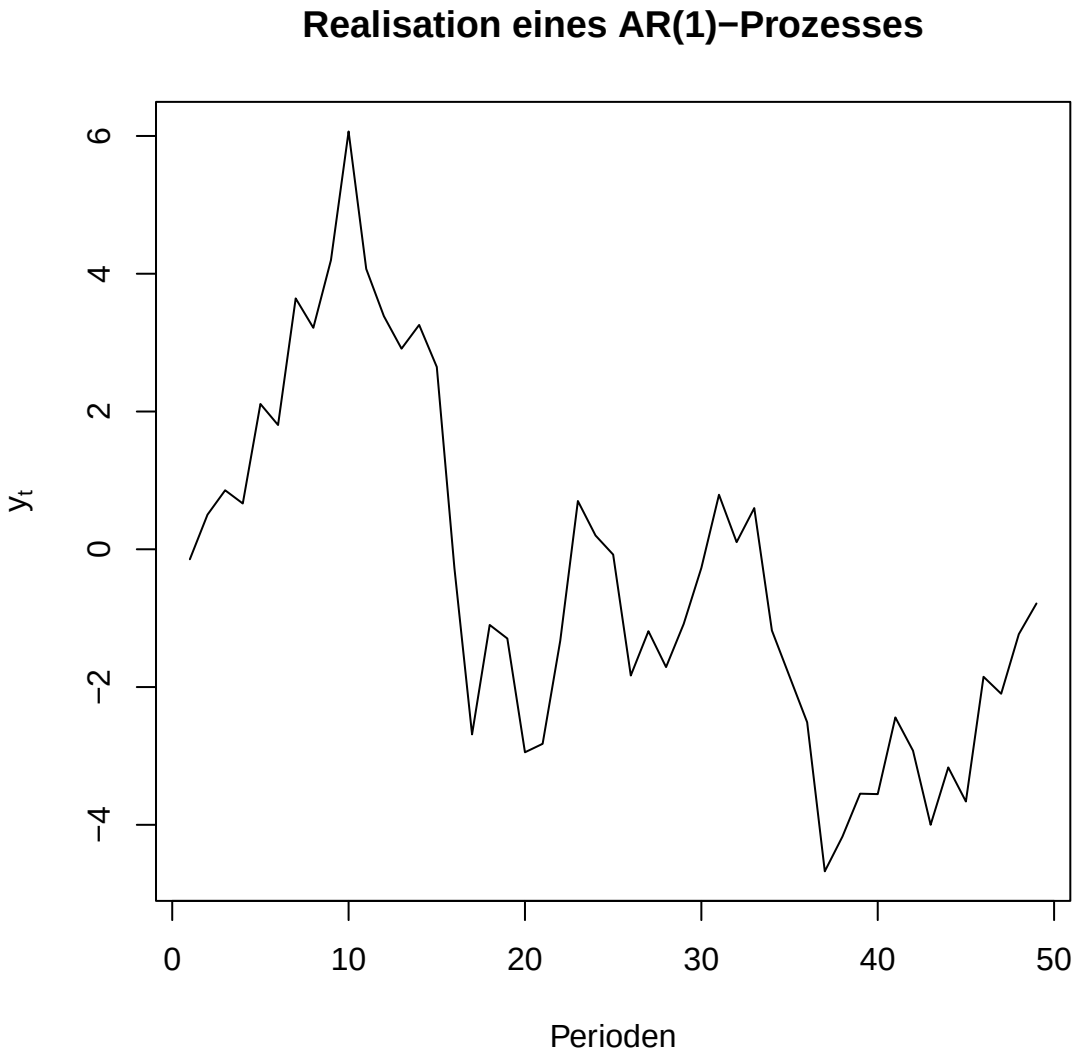

Abbildung 4.1.: Realisation eines AR(1)-Prozesses, siehe Appendix A.9 für das R-Programm

4.3.2. Monte-Carlo-Simulation mehrerer Realisationen

Generiert man insgesamt R Realisationen und zeichnet diese in eine Graphik ein, lassen sich u. a. anschaulich die unbedingten Erwartungswerte eines stochastischen Prozesses illustrieren.

Für die folgende vollständige Spezifikation eines AR(1)-Prozesses

1. $e_t \sim$ i.i.d.$N(0, 1)$,
2. $\nu = 1$ (wenn Konstante berücksichtigt),
3. $\alpha = 0.9$,
4. $y_0 = 5$,

zeigt Abbildung 4.2 50 Realisationen von 100 Beobachtungen jeweils

ohne und mit Konstante, die mit dem R-Programm in Appendix A.10 erzeugt wurden.

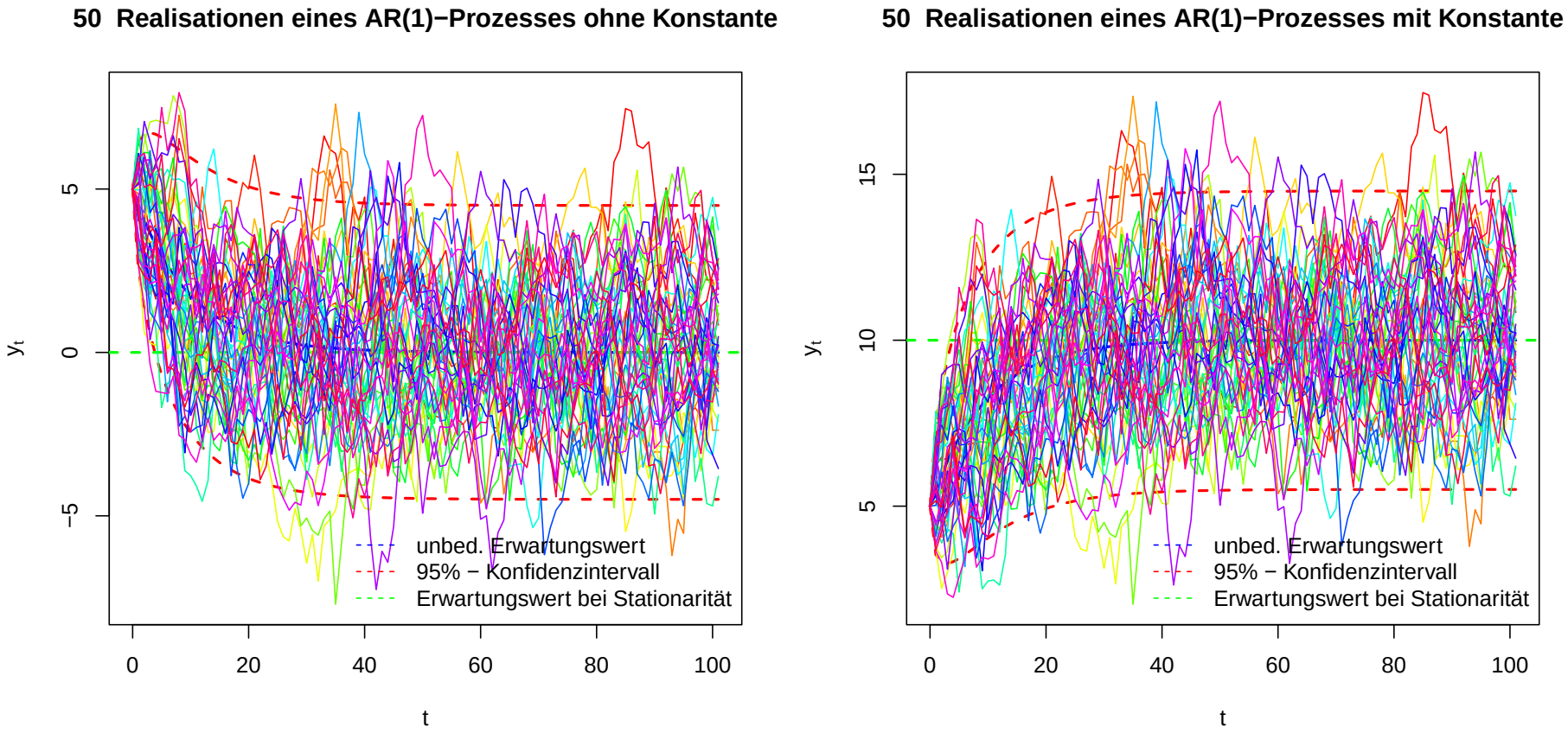

Abbildung 4.2.: 50 Realisationen eines AR(1)-Prozesses ohne und mit Konstante, siehe Appendix A.10 für das R-Programm

In beiden Graphiken zeigt sich deutlich wie sich innerhalb der ersten 50 Perioden der unbedingte Erwartungswert verändert. Diese Perioden werden häufig als Einschwingphase bezeichnet. Danach sind hier Abweichungen des unbedingten Erwartungswertes vom Erwartungswert eines stationären Prozesses vernachlässigbar. Deshalb muss man bei Simulationen von stationären Prozessen zumindest die ersten 50 Perioden löschen, um so die Einschwingphase zu vermeiden.

Es zeigt sich auch, dass nur in wenigen Fällen die Realisationen außerhalb der jeweiligen 95%-Konfidenzintervalle liegen.

4.4. Ausblick: Weitere stochastische Prozesse

- **Autoregressive Prozesse der Ordnung p (AR(p)-Prozesse)**

$$y_t = \nu + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} + e_t, \quad e_t \sim WN(0, \sigma^2).$$

- **Moving Average Prozesse der Ordnung q (MA(q)-Prozesse)**

$$y_t = \mu + e_t + \phi_1 e_{t-1} + \ldots + \phi_q e_{t-q}, \quad e_t \sim WN(0, \sigma^2).$$

- **Autoregressive Moving Average Prozesse (ARMA(p, q-Proz.))**

$$\begin{aligned} y_t = \nu &+ \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} \\ &+ e_t + \phi_1 e_{t-1} + \ldots + \phi_q e_{t-q}, \quad e_t \sim WN(0, \sigma^2). \end{aligned}$$

Die Analyse der jeweiligen statistischen Eigenschaften (Stabilitäts-, Stationaritätsbedingungen, etc.) erfordert z.T. mehr Aufwand, wird nur für AR(p) (später) durchgeführt.

4.5. Schätzung von autoregressiven (AR(p)) Modellen

4.5.1. Notation

- Man erhält das Modell in üblicher Notation:

$$y_t = \mathbf{x}_t\boldsymbol{\beta} + e_t,$$

  indem man mit $k = p$ folgendes verwendet:

$$\mathbf{x}_t = \begin{pmatrix} 1 & x_{t1} & x_{t2} & \cdots & x_{tk} \end{pmatrix} = \begin{pmatrix} 1 & y_{t-1} & \cdots & y_{t-p} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \nu \\ \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix}.$$

- Hieraus lässt sich das Stichprobenmodell leicht in Matrixschreibweise

formulieren:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e},$$

indem man folgende Matrizen definiert:

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix}, \quad \mathbf{e} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_T \end{pmatrix}$$

$$\mathbf{X} = \begin{pmatrix} 1 & y_0 & y_{-1} & \cdots & y_{1-p} \\ 1 & y_1 & y_0 & \cdots & y_{2-p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & y_{T-1} & y_{T-2} & \cdots & y_{T-p} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \nu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{pmatrix}.$$

4.5.2. Schätzung

- Die Schätzung der Parameter von AR(p)-Prozessen kann mittels **OLS** erfolgen.
- Die Schätzung von MA(q)- und ARMA(p, q)-Prozessen ist schwieriger und wird hier nicht behandelt.

4.5.3. Eigenschaften

Berücksichtigung verzögert endogener Variablen führt zur **Verletzung von Annahme TS.3 (strenge Exogenität)**.

- Schätzer sind für jede Stichprobengröße T verzerrt!

- OLS-Schätzer nicht exakt normalverteilt, selbst wenn $e_t \sim N(0, \sigma^2)$.
- Schätzeigenschaften hängen von den autoregressiven Parametern $\alpha_1, \ldots, \alpha_p$ und den deterministischen Termen sowie von der Stichprobengröße T ab. (Illustration in Abschnitt 4.6). Einfachster Fall: AR(p)-Prozess ist stationär.
- Es ist jedoch möglich, die Schätzeigenschaften approximativ zu bestimmen und zwar umso genauer, je mehr Beobachtungen vorliegen. $\longrightarrow$ **asymptotische Schätzeigenschaften**.

- **Nachweis der Parameterverzerrung für ein AR(1)-Modell**

$$y_t = \alpha y_{t-1} + e_t.$$

Der OLS-Schätzer lautet:

$$\hat{\alpha} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \frac{\sum_{t=1}^{T} y_{t-1}y_t}{\sum_{t=1}^{T} y_{t-1}^2} \overset{TS.1}{=} \alpha + \frac{\sum_{t=1}^{T} y_{t-1}e_t}{\sum_{t=1}^{T} y_{t-1}^2} \quad (4.5)$$

Es gilt:

$$E[e_t|\mathbf{X}] = E[e_t|y_0, \ldots, y_{T-1}] \neq 0,$$

denn e_t ist z.B. über y_{t+1} auch in der Bedingung enthalten, da gemäß der "Lösung" eines stationären AR(1)-Prozesses (4.3) $y_{t+1} = \sum_{j=0}^{\infty} \alpha^j e_{t+1-j}$ ist. Damit folgt:

$$E[\hat{\alpha}] \neq \alpha.$$

Also ist TS.3 verletzt, eine notwendige Bedingung für Unverzerrtheit.

- Der OLS-Schätzer ist nicht exakt normalverteilt, da Einsetzen der Moving-Average Darstellung eines AR(1)-Prozesses in (4.5) folgendes ergibt:

$$\hat{\alpha} = \alpha + \frac{\sum_{t=1}^{T} \left(\sum_{j=0}^{\infty} \alpha^j e_{t-1-j} \right) e_t}{\sum_{t=1}^{T} \left(\sum_{j=0}^{\infty} \alpha^j e_{t-1-j} \right)^2}.$$

  Man sieht, dass die Fehler $\{e_t : t = \ldots, -2, -1, 0, 1, \ldots, T\}$ in hochgradig nichtlinearer Weise in den OLS-Schätzer eingehen, so dass $\hat{\alpha}$ nicht (exakt) normalverteilt sein kann.

  Allerdings kann die unbekannte Verteilung durch eine Normalverteilung approximiert werden, *wenn* genügend Beobachtungen T vorliegen (siehe Kapitel 5, asymptotische Eigenschaften).

4.6. Analyse der Schätzeigenschaften für endliche Stichproben mittels Monte-Carlo-Simulationen

Indem man die Generierung eines AR(1)-Prozesses und die entsprechende Parameterschätzung R-mal durchführt, lässt sich der Mittelwert der Schätzungen für die R Realisationen berechnen:

$$\frac{1}{R}\sum_{r=1}^{R}\hat{\alpha}^{(r)}.$$

Hier bezeichnet $\hat{\alpha}^{(r)}$ die Schätzung von α für die r-te Realisation. Je größer die Zahl der Replikationen R, desto geringer wird im Allgemeinen der Approximationsfehler zwischen dem aus der Simulation bestimmten Wert und dem wahren (aber meist unbekannten) Wert.

Zusätzlich kann die Zahl der verwendeten Beobachtungen T variiert

werden. Tabelle 4.1 zeigt die Mittelwerte der $\hat{\alpha}$ für verschiedene T, welche mittels Monte-Carlo-Simulationen mit dem in Appendix A.11 aufgelisteten R-Programm gewonnen wurden. Hierzu wurden für die Stichprobengrößen $T = \{20, 50, 100, 500, 1000, 10000\}$ ein AR(1)-Prozess mit $\alpha = 0.9$ insgesamt $R = 10000$ mal generiert und jeweils α geschätzt.

Abbildung 4.3 zeigt die aus denselben Monte-Carlo-Simulationen resultierenden Verteilungen des geschätzten AR-Parameters α.

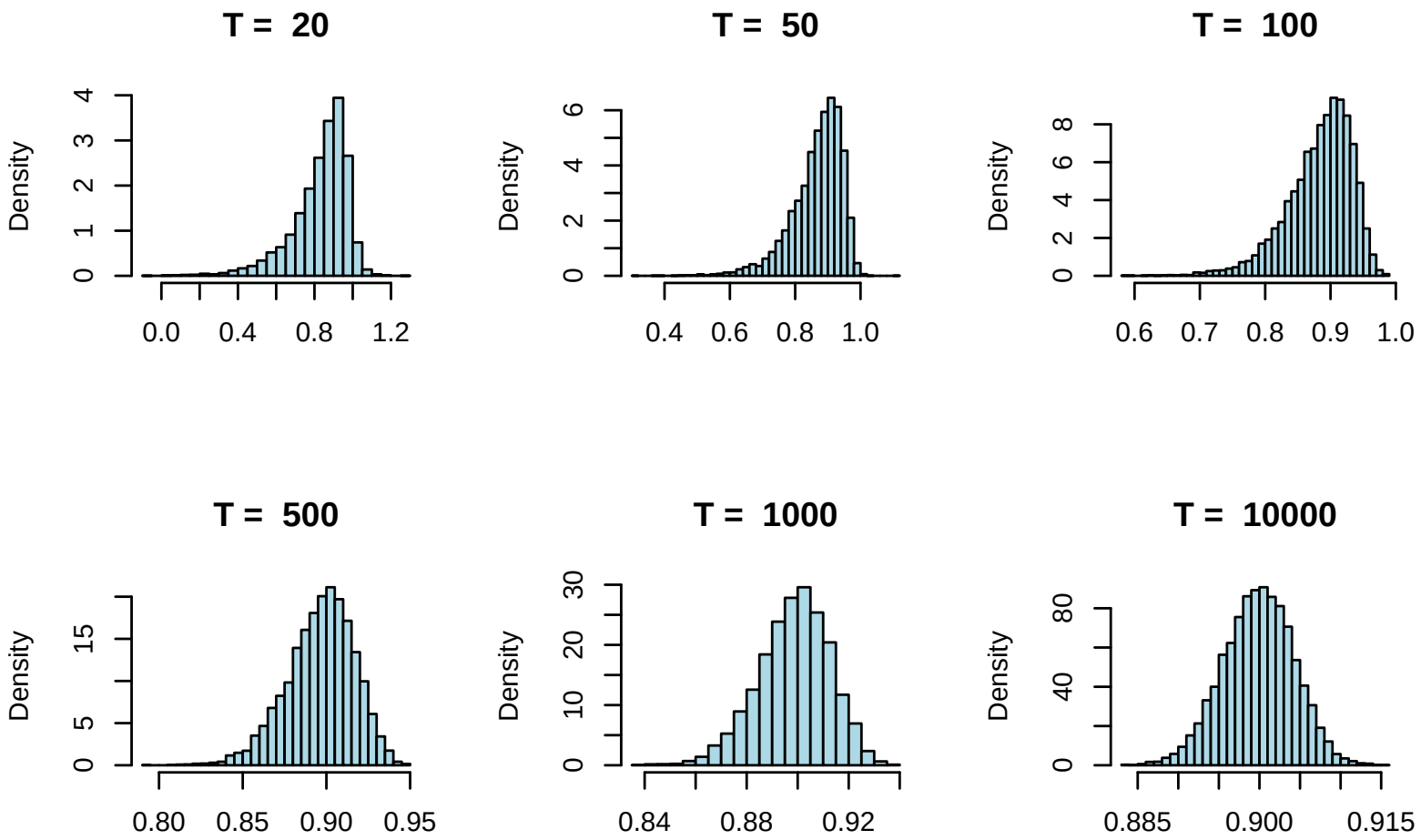

Abbildung 4.3.: Simulierte Verteilungen auf Basis von $R = 10000$ Generierungen und Schätzungen des AR-Parameters $\hat{\alpha}$ eines AR(1)-Prozesses mit $\alpha = 0.9$ gegeben $T = \{20, 50, 100, 500, 1000, 10000\}$ Beobachtungen, siehe Appendix A.11 für das R-Programm

Tabelle 4.1.: Mittelwerte und Standardabweichungen für $\hat{\alpha}$ auf Basis von $R = 10000$ Generierungen und Schätzungen des AR-Parameters eines AR(1)-Prozesses mit $\alpha = 0.9$ und T Beobachtungen, siehe Appendix A.11 für das R-Programm

	Mittelwerte	Standardabweichungen
T = 20	0.831027	0.145496
T = 50	0.866797	0.077428
T = 100	0.883125	0.049866
T = 500	0.896511	0.019978
T = 1000	0.898531	0.013920
T = 10000	0.899856	0.004345

Ergebnisse

- Gemäß Tabelle 4.1 beträgt die Abweichung der gemittelten geschätzen Werte vom wahren Wert -0.068973 für $T = 20$, dagegen lediglich -0.000144 für $T = 2000$. Ebenso reduziert sich Standardabweichung.

- Beides zusammen suggeriert Konsistenz des OLS-Schätzers (=asymptotische Eigenschaft).

- In Abbildung 4.3 ist für $T = 20$ die Verteilung von $\hat{\alpha}$ linksschief (=rechtssteil) und damit nicht normal. Für $T = 2000$ ist keine Schiefe mehr zu sehen.

- Dies suggeriert die Normalverteilung als gute Approximation (=asymptotische Eigenschaft).

Zu lesen: Section 11.1 in Wooldridge (2009).

5. Asymptotische Eigenschaften des OLS-Schätzers: Schätzeigenschaften des OLS-Schätzers bei Verletzung der klassischen OLS-Annahmen

5.1. Grundlagen

- Erinnerung: Der OLS-Schätzer $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ ist ein Vektor von Zufallsvariablen, solange noch keine Stichprobe gezogen wurde.

- Man stelle sich nun vor, man würde eine Stichprobe der Länge T

ziehen und $\hat{\boldsymbol{\beta}}$ berechnen. Um zu verdeutlichen, dass genau T Beobachtungen verwendet wurden, indizieren wir den $((k+1)\times 1)$-Schätzvektor mit T und schreiben $\hat{\boldsymbol{\beta}}_T$.

Was wird passieren, wenn man die Stichprobengröße von T auf $T+1$ erhöhen könnte:

$$\hat{\boldsymbol{\beta}}_{T+1} = \hat{\boldsymbol{\beta}}_T \quad \text{oder} \quad \hat{\boldsymbol{\beta}}_{T+1} \neq \hat{\boldsymbol{\beta}}_T?$$

Hierüber kann man im Allgemeinen nichts aussagen, aber:

- Falls man nichts über eine einzelne Stichprobe aussagen kann, kann man etwas über die **Veränderung der Wahrscheinlichkeitsverteilungen bzw. deren Eigenschaften wie Erwartungswert oder Varianz in Abhängigkeit von T aussagen**?

Damit könnte man Wahrscheinlichkeitsaussagen für eine noch zu ziehende Stichprobe machen!

- Erinnerung: Was passierte bei Monte Carlo-Simulationen zur Untersuchung von Schätzeigenschaften des OLS-Schätzers eines AR(1)-Modells in Abschnitt 4.6?
  - $E[\hat{\alpha}_{20}]$ war weiter weg von wahrem Wert als $E[\hat{\alpha}_{2000}]$.
  - $Var(\hat{\alpha}_{20})$ größer als $Var(\hat{\alpha}_{2000})$.
  - Wahrscheinlichkeitsverteilung von $\hat{\alpha}_{20}$ schiefer als Wahrscheinlichkeitsverteilung von $\hat{\alpha}_{2000}$. Letztere schaut "normaler" aus.

  Was passiert jeweils für $T = 200000$? Antwort durch Monte Carlo Simulationen oder durch Anwendung der Theorie mathematischer

Statistik. Letzteres folgt jetzt.

- Betrachtet man nun alle möglichen Stichprobenlängen $T = 1, 2, \ldots$, dann ist $\hat{\alpha}_T$ eine **Folge von Zufallszahlen** und $\hat{\boldsymbol{\beta}}_T$ eine **Folge von Zufallsvektoren**.

- In der **mathematischen Statistik** hat man ein Instrumentarium geschaffen, um das Verhalten von Folgen von Zufallsvariablen allgemein zu charakterisieren und zu analysieren.

5.1.1. Definition: Konvergenz in Wahrscheinlichkeit

Gegeben sei eine Folge von Zufallszahlen $\theta_T, T = 1, 2, 3, \ldots$. Dann konvergiert θ_T in Wahrscheinlichkeit gegen die Zufallsvariable θ, wenn für jedes $\varepsilon > 0$ gilt:

$$\lim_{T \to \infty} P\left(|\theta_T - \theta| < \varepsilon\right) = 1.$$

Äquivalent lässt sich schreiben: wenn es für beliebig kleine $\varepsilon > 0$ und für beliebig kleine $\delta > 0$ ein T_0 gibt, so dass für alle $T > T_0$ gilt:

$$P\left(|\theta_T - \theta| < \varepsilon\right) > 1 - \delta.$$

Kurzschreibweisen: $\theta_T \xrightarrow{P} \theta$ oder $\operatorname{plim}_{T \to \infty} \theta_T = \theta$. Man bezeichnet dann θ als **Wahrscheinlichkeitsgrenzwert** (**probability limit**).

Zeichnen Sie die Wahrscheinlichkeitsverteilungen von θ_T für verschiedene T, um Konvergenz in Wahrscheinlichkeit zu illustrieren!

5.1.2. Definition: Konvergenz in Verteilung

Sei θ_T eine Folge von Zufallsvariablen ($T = 1, 2, 3, \ldots$) mit Verteilungsfunktionen F_T. Sei θ eine Zufallsvariable mit Verteilungsfunktion F. θ_T konvergiert in Verteilung gegen θ, wenn für alle x gilt:

$$\lim_{T \longrightarrow \infty} F_T(x) = F(x) \quad \text{bzw.} \quad \lim_{T \longrightarrow \infty} P(\theta_T \leq x) = P(\theta \leq x).$$

Kurzschreibweise: $\theta_T \xrightarrow{d} \theta$.

Der bekannteste Fall ist Konvergenz gegen eine Standardnormalverteilung:

$$P(\theta_T \leq x) \longrightarrow \Phi(x),$$

wobei $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung bezeichnet. Man sagt dann, dass θ_T eine **asymptotische Standardnormalverteilung** aufweist.

5.1.3. Einfaches Beispiel

Es seien X und Y jeweils standardnormalverteilte Zufallsvariablen. Gegeben sei die Folge von Zufallsvariablen $X_T = X + \frac{1}{T}Y$, $T = 1, 2, \ldots$.

X_T konvergiert in Wahrscheinlichkeit gegen X, also $X_T \xrightarrow{P} X$ bzw. $\operatorname{plim}_{T\to\infty} X_T = X$, da

$$\lim_{T\to\infty} P(|X_T - X| < \varepsilon) = \lim_{T\to\infty} P\left(\left|\frac{1}{T}Y\right| < \varepsilon\right) = 1.$$

Dies gilt, da mit zunehmendem T immer mehr mögliche Realisationen von $\frac{1}{T}Y$ im Intervall $(-\varepsilon, \varepsilon)$ liegen. Konvergenz in Wahrsch.keit setzt also voraus, dass die Varianz der Differenz $|X_t - X|$ gegen Null geht.

X_T konvergiert in Verteilung gegen X *oder* Y, ist also asymptotisch standardnormalverteilt:

$$\lim_{T\to\infty} P(X_T < a) = P\left(X < a\right) = P(Y < a) = \Phi(a).$$

5.1.4. Weiteres einfaches Beispiel

Es seien X und Y unabhängige, jeweils standardnormalverteilte Zufallsvariablen. Gegeben sei für $T = 1, 2, \ldots$ die Folge von Zufallsvariablen:

$$Z_T = \frac{X}{\sqrt{2}} + (-1)^T \frac{Y}{\sqrt{2}}.$$

Eigenschaften von Z_T: Egal wie groß T ist, es gilt:

$$Z_T = \begin{cases} \frac{1}{\sqrt{2}}X + \frac{1}{\sqrt{2}}Y & \text{falls } T \text{ gerade,} \\ \frac{1}{\sqrt{2}}X - \frac{1}{\sqrt{2}}Y & \text{falls } T \text{ ungerade.} \end{cases}$$

Damit kann Z_T weder gegen X, noch gegen Y konvergieren, aber sowohl die Summe als auch die Differenz sind normalverteilt mit Mittelwert 0 und Varianz 1. Warum?

Deshalb konvergiert Z_T in Verteilung gegen die Standardnormalverteilung, also:

$$\lim_{T\to\infty} P(Z_T < a) = P\left(X < a\right) = P(Y < a) = \Phi(a).$$

Fazit

Konvergenz in Wahrscheinlichkeit impliziert Konvergenz in Verteilung, aber nicht umgekehrt. (Dies lässt sich allgemein zeigen.)

5.1.5. Regeln für plim's

- **Slutzky's Theorem**: Falls $\operatorname{plim}\theta_T = \theta$ und $g(\cdot)$ an der Stelle θ stetig ist, ist $\operatorname{plim} g(\theta_T) = g(\theta)$.

- Seien $\{\mathbf{x}_T\}$ und $\{\mathbf{y}_T\}$ Folgen von $(K \times 1)$-Vektoren und $\{\mathbf{A}_T\}$ eine Folge von $(K \times K)$-Matrizen. Falls $\operatorname{plim} \mathbf{x}_T$, $\operatorname{plim} \mathbf{y}_T$ und $\operatorname{plim} \mathbf{A}_T$ existieren, gilt:

  - $\operatorname{plim}(\mathbf{x}_T \pm \mathbf{y}_T) = \operatorname{plim} \mathbf{x}_T \pm \operatorname{plim} \mathbf{y}_T$,
  - $\operatorname{plim}(\mathbf{x}_T{}' \mathbf{y}_T) = (\operatorname{plim} \mathbf{x}_T)'(\operatorname{plim} \mathbf{y}_T)$,
  - $\operatorname{plim}(\mathbf{A}_T \mathbf{y}_T) = (\operatorname{plim} \mathbf{A}_T)(\operatorname{plim} \mathbf{y}_T)$,
  - Falls $\mathbf{x}_T \xrightarrow{d} \mathbf{x}$ und $\operatorname{plim} \mathbf{A}_T = \mathbf{A}$, dann gilt $\mathbf{A}_T \mathbf{x}_T \xrightarrow{d} \mathbf{A}\mathbf{x}$.

  (Siehe Appendix C.3, Property PLIM.1 und PLIM.2 in Wooldridge (2009) für den skalaren Fall $K = 1$ oder z.B. Proposition C.2, S. 683 in Lütkepohl (2005) für den Vektorfall.)

5.2. Anwendung auf Schätzer

Es sei $y_t \sim$ i.i.d.$N(\mu, \sigma^2)$, $t = 1, 2, \ldots, T$.

$$\hat{\mu}_T = \frac{y_1}{T} + \frac{y_2}{T} + \cdots + \frac{y_T}{T} = \frac{1}{T} \sum_{t=1}^{T} y_t$$

ist dann der OLS-Schätzer für den Mittelwert μ einer Zufallsstichprobe $\{y_t : t = 1, 2, \ldots, T\}$.

Gemäß den Eigenschaften von Summen unabhängiger normalverteilter Zufallsvariablen gilt für jedes beliebige T:

$$\hat{\mu}_T \sim N\left(\mu, \frac{1}{T}\sigma^2\right)$$

und

$$\sqrt{T}\,(\hat{\mu}_T - \mu) \sim N(0, \sigma^2) \quad \text{bzw.}$$
$$\frac{\hat{\mu}_T - \mu}{\sigma/\sqrt{T}} \sim N(0,1). \tag{5.1}$$

Zeigen Sie das jeweils!

Was passiert, wenn y_t i.i.d. $\chi^2(1)$-verteilt ist?

Oder einer anderen beliebigen Verteilung folgt?

5.2.1. Konvergenz in Wahrscheinlichkeit

- **Gesetz der großen Zahlen** (Law of Large Numbers, LLN)
  Es sei $y_t \sim$ i.i.d. mit Mittelwert $|\mu| < \infty$, $t = 1, 2, \ldots, T$. Dann gilt für den Mittelwertschätzer einer Zufallsstichprobe $\{y_t\}_{t=1}^{T}$,

$$\hat{\mu}_T = \frac{1}{T} \sum_{t=1}^{T} y_t,$$

  dass

$$\operatorname*{plim}_{T \to \infty} \hat{\mu}_T = \mu.$$

- Ein Schätzer für den Parametervektor $\boldsymbol{\beta}$ heißt **konsistent**, wenn $\operatorname{plim}_{T \to \infty} \hat{\boldsymbol{\beta}}_T = \boldsymbol{\beta}$. Entsprechend heißt ein Schätzer für den Fehler u_t konsistent, wenn gilt $\operatorname{plim}_{T \to \infty} \hat{u}_t = u_t$. Beachte: Im ersten Fall ist der Grenzwert nicht-stochastisch, im zweiten Fall stochastisch.

Konsistenz impliziert, dass

- der Schätzer **asymptotisch unverzerrt** ist (d.h. für $T \longrightarrow \infty$)
- und dass die Varianz des Schätzers asymptotisch gegen Null geht (für $T \longrightarrow \infty$ konzentriert sich $\hat{\boldsymbol{\beta}}_T$ immer mehr um $\boldsymbol{\beta}$).

Damit ist der Mittelwertschätzer $\hat{\mu}_T$ konsistent.

- Konsistenz ist eine Minimalanforderung an einen brauchbaren Schätzer.
- Ist ein Schätzer konsistent, sind auch nichtlineare Transformationen des Schätzers konsistent, z.B. $\frac{\hat{\beta}_1}{\hat{\beta}_2}$ oder $\hat{\beta}_1 \hat{\beta}_2$. Warum?
- Es gibt Versionen des Gesetzes der Großen Zahlen, die gelten, wenn die i.i.d. Annahme durch die Annahme schwacher Abhängigkeit ("weak dependence") (plus weitere technische Bedingungen) ersetzt wird.

5.2.2. Konvergenz in Verteilung

- **Zentraler Grenzwertsatz** (Central Limit Theorem (CLT))
  Es sei $y_t \sim$ i.i.d.(μ, σ^2), $t = 1, 2, \ldots, T$, $|\mu| < \infty$, $\sigma^2 < \infty$, eine Zufallsstichprobe. Dann gilt für den Mittelwertschätzer

$$\hat{\mu}_T = \frac{1}{T} \sum_{t=1}^{T} y_t,$$

  dass

$$\sqrt{T}\,(\hat{\mu}_T - \mu) \xrightarrow{d} N(0, \sigma^2), \quad \frac{\hat{\mu}_T - \mu}{\sigma/\sqrt{T}} \xrightarrow{d} z \sim N(0, 1), \quad (5.2)$$

  bzw. in alternativer Darstellung

$$P\left(\frac{\hat{\mu}_T - \mu}{\sigma/\sqrt{T}} < x\right) \longrightarrow \Phi(x).$$

- Sind die y_t nicht mehr normalverteilt, wird in (5.1) "$\sim$" durch "$\xrightarrow{d}$" ersetzt und die Normalverteilung gilt nur noch asymptotisch, also "in großen Stichproben".

- Aus (5.2) folgt nicht, dass nichtlineare Transformationen, z.B. $\hat{\mu}^2$, ebenfalls asymptotisch (standard)normalverteilt sind!! Warum?

- Es gibt auch zentrale Grenwertsätze, die gelten, wenn die i.i.d. Annahme durch die Annahme schwacher Abhängigkeit ("weak dependence") (plus weitere technische Bedingungen) ersetzt wird.

Zu lesen: Appendix C.3 in Wooldridge (2009).

5.3. Asymptotische Eigenschaften von OLS bei Zeitreihendaten und nichtnormalverteilten Fehlern

Betrachtet wird das Regressionsmodell

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \quad \text{bzw.} \quad y_t = \mathbf{x}_t\boldsymbol{\beta} + u_t, \quad t = 1, 2, \ldots, T,$$

und der OLS-Schätzer für eine Stichprobe der Länge T,

$$\hat{\boldsymbol{\beta}}_T = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

5.3.1. Annahmen

- **Annahme TS.1'** (Linearität in den Parametern und schwache Abhängigkeit)
  Es gilt Annahme TS.1 (siehe Abschnitt 2.2.1). Zusätzlich ist der

gemeinsame Zeitreihenprozess $\{(y_t, \mathbf{x}_t) : t = 1, 2, \ldots\}$ streng stationär und schwach abhängig. Darüber hinaus sind alle (hier nicht weiter spezifierten) Regularitätsannahmen erfüllt, so dass ein Gesetz der Großen Zahlen und ein Zentraler Grenzwertsatz (jeweils für nicht-i.i.d.-Beobachtungen) angewendet werden kann.

- **Annahme TS.2'** (Keine perfekte Kollinearität)
  Wie TS.2 (siehe Abschnitt 2.2.1).

- **Annahme TS.3'** (Bedingter Erwartungswert Null)
  Abschwächung von TS.3 (siehe Abschnitt 2.2.1): Der Fehler u_t ist von allen **gegenwärtigen** erklärenden Variablen $\mathbf{x}_t$ unabhängig:

$$E[u_t|\mathbf{x}_t] = 0, \quad t = 1, 2, \ldots.$$

  Man spricht von **vorherbestimmten Regressoren**, **partieller Unabhängigkeit** oder **"contemporaneous exogeneity"**.

- **Annahme TS.4'** (Bedingte Homoskedastie)
  Schwächt Annahme TS.4 (siehe Abschnitt 2.2.2) ab. Die Fehler müssen nur bedingt auf die **gegenwärtigen** erklärenden Variablen $\mathbf{x}_t$ homoskedastisch sein:

$$Var(u_t|\mathbf{x}_t) = \sigma^2.$$

- **Annahme TS.5'** (Keine Autokorrelation in den Fehlern, No Serial Correlation)
  Schwächt Annahme TS.5 (siehe Abschnitt 2.2.2) ab. Für alle $t \neq s$ gilt: $Cov[u_t, u_s|\mathbf{x}_t, \mathbf{x}_s] = 0$.

- Beachte:
  - Annahme TS.6 (Normalverteilte Fehler) kann entfallen.
  - Die Annahme strenger Stationarität in TS.1' kann zu schwacher

Stationarität abgeschwächt werden, falls zusätzliche Annahmen über den Prozess der Fehler getroffen werden. Siehe z.B. MA-Veranstaltung **Methoden der Ökonometrie**, Abschnitt **Dynamische lineare Regressionsmodelle**.

5.3.2. Schätzeigenschaften

- **Theorem 5.1**: **Konsistenz des OLS-Schätzers**
  Unter TS.1', TS.2' und TS.3' ist der OLS-Schätzer konsistent, d.h.

$$\text{plim}_{T\to\infty}\,\hat{\boldsymbol{\beta}}_T = \boldsymbol{\beta}.$$

  (Entspricht Theorem 11.1 in Wooldridge (2009).)

- **Theorem 5.2**: **Asymptotische Normalverteilung**
  Unter TS.1' bis TS.5' ist der OLS-Schätzer asymptotisch normal-

verteilt:

$$\sqrt{T}\left(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}\right) \xrightarrow{d} N\left(0, \sigma^2 \mathbf{A}^{-1}\right),$$

mit $\operatorname{plim}_{T\to\infty} \frac{1}{T}\mathbf{X}'\mathbf{X} = \mathbf{A}$, wobei $\mathbf{A}$ unter den gegebenen Annahmen invertierbar ist.

(Entspricht Theorem 11.2 in Wooldridge (2009).)

5.3.3. Diskussion

In der Praxis ist $\mathbf{A}$ unbekannt und wird durch $\frac{1}{T}\mathbf{X}'\mathbf{X}$ geschätzt. Damit ergibt sich die approximative Normalverteilung:

$$\sqrt{T}\left(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}\right) \approx N\left(0, \sigma^2\left(\frac{1}{T}\mathbf{X}'\mathbf{X}\right)^{-1}\right)$$

bzw. nach Herausmultiplizieren von T:

$$\hat{\boldsymbol{\beta}}_T \approx N\left(\boldsymbol{\beta}, \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}\right),$$

so dass im Vergleich zum klassischen normalen Regressionsmodell nunmehr die Normalverteilung nur noch approximativ gilt, jedoch die gleiche Formel verwendet werden kann.
Dabei — hier nicht mehr sichtbar — reduziert sich der Approximationsfehler mit dem Stichprobenumfang T.

Für eine gegebene vollständige Modellspezifikation kann der Approximationsfehler der asymptotischen Normalverteilung mit Hilfe von Monte Carlo Simulationen bestimmt werden, wie beispielsweise für spezifische AR(1)-Prozesse in Abschnitt 4.6.

5.3.4. Wann ist die Annahme TS.1' erfüllt?

- **Klassisches statisches Regressionsmodell**

$$y_t = \mathbf{x}_t\boldsymbol{\beta} + u_t, \quad u_t \sim \text{i.i.d.}(0, \sigma^2),$$

wobei die Annahmen MLR.1 bis MLR.5 erfüllt sind, jedoch die Normalverteilungsannahme MLR.6 nicht.

Man beachte, dass i.i.d.-verteilte Stichprobenbeobachtungen $\{(y_t, \mathbf{x}_t) : t = 1, 2, \ldots, T\}$ (Zufallsstichprobe) trivialerweise streng stationär

sind und keine Abhängigkeit aufweisen, so dass durch die Annahmen MLR.1 und MLR.2 dieser Teil der Annahme TS.1' erfüllt ist.

Beachte: TS.1' deckt nicht den Fall von Regressoren mit deterministischen Trends ab, da trendbehaftete Regressoren nicht stationär sind.

- **Finite Distributed Lag Modell** (siehe Abschnitte 2.1 und 2.2)

$$y_t = \alpha_0 + \delta_0 z_t + \delta_1 z_{t-1} + \cdots + \delta_q z_{t-q} + \beta_1 x_t + u_t, \quad u_t \sim \text{i.i.d.}(0, \sigma^2).$$

Im Gegensatz zum klassischen Regressionsmodell für Zufallsstichproben, sind, falls die Normalverteilungsannahme TS.6 verletzt ist, die Annahmen TS.1 bis TS.5 für die asymptotische Normalverteilung des OLS-Schätzers nicht hinreichend. Der Grund ist: In TS.1 bis TS.5 wird nichts über die stochastische Abhängigkeit der Regressoren gefordert (TS-Annahmen siehe Abschnitt 2.2).

Also: Sind die Fehler nicht normalverteilt, muss statt der Annahme TS.1 die Annahme TS.1' erfüllt sein, damit die genannte asymptotische Normalverteilung gilt. Siehe hierzu unten den Punkt zu dynamischen Regressionsmodellen. Beachte, dass TS.2' bis TS.5' gleich oder schwächer wie die entsprechenden Annahmen im FDL-Modell sind.

- **Stationäre autoregressive Prozesse der Ordnung** p

$$y_t = \nu + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} + e_t, \quad e_t \sim \text{i.i.d.}(0, \sigma^2).$$

Homoskedastische streng stationäre AR(p)-Prozesse sind schwach abhängig und es existiert jeweils sowohl ein entsprechendes Gesetz der Großen Zahlen als auch ein Zentraler Grenzwertsatz, so dass TS.1' erfüllt ist. Genaue Bedingungen für Stationarität werden im MA-Kurs **Applied Financial Econometrics**, Abschnitt 2.1 oder z.B. in Kirchgässner und Wolters (2008, Abschnitt 2.1) oder Neusser

und Wagner (2022, Abschnitt 2.3) besprochen. Ist die Ordnung von AR-Prozessen korrekt spezifiziert, ist auch Annahme TS.3' erfüllt.

- **Dynamische Regressionsmodelle** (vgl. auch Abschnitt 7.1)

$$\begin{aligned} y_t = \nu &+ \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} \\ &+ \delta_0 z_t + \delta_1 z_{t-1} + \cdots + \delta_q z_{t-q} \\ &+ \beta_1 x_{t1} + \cdots + \beta_l x_{tl} + u_t, \quad u_t \sim \text{i.i.d.}(0, \sigma^2). \end{aligned}$$

TS.1' ist erfüllt, wenn der enthaltene AR(p)-Prozess streng stationär ist und

  - die exogenen Regressoren i.i.d. Variablen sind. Dann ist $\{y_t\}$ streng stationär und schwach abhängig, wenn die Verteilung der Startwerte der streng stationären Verteilung entspricht;

  - die exogenen Regressoren selbst einem (eigenen) stochastischen

Prozess folgen, der streng stationär und schwach abhängig ist und der Prozess in der streng stationären Verteilung startet.

Siehe auch Kapitel 6.

Noch wichtig: Die Annahme strenger Stationarität kann zur Annahme schwacher Stationarität abgeschwächt werden, siehe hierzu z. B. Neusser und Wagner (2022, Kapitel 2.3) oder die Folien zum Mastermodul Applied Financial Econometrics.

5.3.5. Beweisskizzen

Theorem 5.1: Konsistenz des OLS-Schätzers

- Siehe Appendix E.4 in Wooldridge (2009) für einen allgemeinen Beweis von Theorem 5.1.
- **Beweisskizze** für das einfache Regressionsmodell im Fall einer Zufallsstichprobe (y_t, x_t) (siehe auch Appendix 5A, S. 182, in Wooldridge (2009)):
  - Der OLS-Schätzer für β_1 im einfachen Regressionsmodell lautet (siehe **Einführung in die Ökonometrie**):

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{T}(x_t - \bar{x})(y_t - \bar{y})}{\sum_{t=1}^{T}(x_t - \bar{x})^2}.$$

Einsetzen des korrekten Modells und Dividieren von Zähler und Nenner durch T ergibt:

$$\hat{\beta}_1 = \beta_1 + \frac{\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})u_t}{\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^2}. \tag{5.3}$$

– Zähler- und Nennerterm sind Mittelwertschätzer der Zufallsvariablen $(x_t - \bar{x})u_t$ bzw. $(x_t - \bar{x})^2$.

– Man beachte, dass die Zufallsvariablen $(x_t - \bar{x})u_t$ sowie $(x_t - \bar{x})^2$, $t = 1, 2, \ldots$, nicht i.i.d. sind, so dass jeweils für den Zähler- und den Nennerterm nicht das Gesetz der Großen Zahlen aus Abschnitt 5.2.1 angewendet werden kann.

Dies Problem lässt sich lösen, indem man $x_t - \bar{x}$ geeignet erweitert durch:

$$x_t - \bar{x} = \underbrace{x_t - \mu}_{\text{i.i.d. Zufallsvariable}} + \mu - \bar{x}.$$

Im Folgenden bezeichne $\sigma_x^2 = Var(x)$.

– Für den Zählerterm erhält man:

$$\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})u_t = \underbrace{\frac{1}{T}\sum_{t=1}^{T}(x_t - \mu)u_t}_{\text{LLN anwendbar}} + \underbrace{\frac{1}{T}\sum_{t=1}^{T}(\mu - \bar{x})u_t}_{(\mu-\bar{x})\frac{1}{T}\sum_{t=1}^{T}u_t}. \quad (5.4)$$

∗ Um den plim des ersten Terms auf der rechten Seite zu berechnen, muss man den Erwartungswert von $(x_t - \mu)u_t$ bestimmen.

Wegen TS.3' und dem Gesetz der iterierten Erwartung (vgl.

Einführung in die Ökonometrie oder Property CE.4 in Appendix B.4 in Wooldridge (2009)) gilt für den Mittelwert:

$$E[(x_t - \mu)u_t] = E\left[(x_t - \mu)E[u_t|x_t]\right] = E[(x_t - \mu) \cdot 0] = 0.$$

Man beachte außerdem, dass der Mittelwert hier gerade einer Kovarianz entspricht, denn $E[(x_t-\mu)u_t] = Cov(x_t, u_t)$ wegen $E[u_t] = 0$. Damit gilt nach dem Gesetz der Großen Zahlen (für i.i.d. Variablen):

$$\text{plim}_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}(x_t-\mu)u_t = E[(x_1-\mu)u_1] = Cov(x_1, u_1) = 0.$$

$*$ Um den plim des zweiten Terms auf der rechten Seite zu berechnen, ist zu bedenken, dass sowohl für den Mittelwertschätzer $\bar{x}$ von x_t als auch für den Mittelwertschätzer von u_t das Gesetz

der Großen Zahlen aus Abschnitt 5.2.1 anwendbar ist:

$$\operatorname*{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x_t = \operatorname*{plim}_{T\to\infty} \bar{x} = \mu,$$

$$\operatorname*{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} u_t = 0.$$

Man erhält also unter der Anwendung der plim-Regeln

$$\operatorname*{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} (x_t - \bar{x})u_t = 0 + 0 \cdot 0 = 0.$$

– Für den Nennerterm erhält man bei entsprechendem Vorgehen:

$$\operatorname*{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^2 = \sigma_x^2,$$

da $E\left[(x-\mu)^2\right] = Var(x) = \sigma_x^2$.

– Angewandt auf (5.3) ergibt sich mit den Regeln für plim's (wobei $T \to \infty$ weggelassen wurde):

$$\operatorname{plim}\hat{\beta}_1 = \beta_1 + \frac{Cov(x,u)}{Var(x)} = \beta_1 + \frac{0}{\sigma_x^2} = \beta_1.$$

- **Anmerkungen**
  - Die obige Vorgehensweise bleibt bei multiplen Regressionsmodellen erhalten. Dann müssen für die jeweiligen Regeln lediglich die entsprechenden Versionen für Vektoren und Matrizen verwendet werden.
  - Liegt keine Zufallsstichprobe vor, z.B. bei einer Stichprobe, die durch einen autoregressiven Prozess generiert wurde, dann bleibt das obige Vorgehen gültig, wenn Annahme TS.1' gilt, d.h. die Stichprobenbeobachtungen streng stationär und schwach abhängig sind und alle technischen Voraussetzungen erfüllt sind, so dass weiter ein Gesetz der Großen Zahlen und ein Zentraler Grenzwertsatz gelten. Siehe Appendix E.4 in Wooldridge (2009) für einen allgemeinen Beweis von Theorem 5.1.

Theorem 5.2: Asymptotische Normalverteilung

- Ein Beweis findet sich in Appendix E.4 in Wooldridge (2009).
- **Beweisskizze** für das einfache Regressionsmodell im Fall einer Zufallsstichprobe (y_t, x_t) (siehe auch Appendix 5A in Wooldridge (2009)):
  - Für die Varianz des ersten Terms auf rechten Seite in (5.4) gilt: wegen TS.3' bis TS.5' und dem Gesetz der iterierten Erwartung (vgl. **Einführung in die Ökonometrie** oder Property CE.4 in Appendix B.4 in Wooldridge (2009)):

$$\begin{aligned} Var\left((x_t - \mu)u_t\right) &= E\left[(x_t - \mu)^2 E[u_t^2|x_t]\right] \\ &= \sigma^2 E\left[(x_t - \mu)^2\right] = \sigma^2\sigma_x^2 \end{aligned}$$

und somit:

$$Var\left(\frac{1}{T}\sum_{t=1}^{T}(x_t-\mu)u_t\right)=\frac{1}{T^2}\sum_{t=1}^{T}Var\left((x_t-\mu)u_t\right)=\frac{1}{T}\sigma^2\sigma_x^2,$$

wobei das erste "=" gilt, weil $Cov((x_t-\mu)u_t,(x_s-\mu)u_s)=0$ für alle $t\neq s$ (Annahme Zufallsstichprobe). Damit gilt:

$$\lim_{T\to\infty}Var\left(\frac{1}{T}\sum_{t=1}^{T}(x_t-\mu)u_t\right)=0.$$

Multipliziert man den ersten Term auf rechten Seite in (5.4) mit $\sqrt{T}$, konvergiert dessen Varianz gegen die Konstante $\sigma^2\sigma_x^2$, da die Division durch T vermieden wird:

$$\lim_{T\to\infty}Var\left(\sqrt{T}\frac{1}{T}\sum_{t=1}^{T}(x_t-\mu)u_t\right)=\sigma^2\sigma_x^2.$$

Damit erfüllt dieser Term die Bedingungen für den Zentralen Grenzwertsatzes (siehe Abschnitt 5.2.2) und es gilt:

$$\sqrt{T}\frac{1}{T}\sum_{t=1}^{T}(x_t-\mu)u_t \xrightarrow{d} W, \quad W \sim N\left(0, \sigma^2\sigma_x^2\right).$$

– Multipliziert man den gesamten Zählerterm mit $\sqrt{T}$ erhält man:

$$\sqrt{T}\frac{1}{T}\sum_{t=1}^{T}(x_t-\bar{x})u_t = \underbrace{\frac{\sqrt{T}}{T}\sum_{t=1}^{T}(x_t-\mu)u_t}_{\text{CLT anwendbar}} + \underbrace{\frac{\sqrt{T}}{T}\sum_{t=1}^{T}(\mu-\bar{x})u_t}_{(\mu-\bar{x})\frac{\sqrt{T}}{T}\sum_{t=1}^{T}u_t}.$$

Betrachtet man das Konvergenzverhalten des zweiten Terms auf

der rechten Seite gilt:

$$(\mu - \bar{x}) \xrightarrow{P} 0,$$
$$\frac{\sqrt{T}}{T} \sum_{t=1}^{T} u_t \xrightarrow{d} V \sim N(0, \sigma^2),$$

so dass insgesamt gemäß den Rechenregeln für plim's gilt:

$$(\mu - \bar{x}) \frac{\sqrt{T}}{T} \sum_{t=1}^{T} u_t \xrightarrow{d} 0 \cdot V = 0$$

und der zweite Term asymptotisch keine Rolle spielt.

– Das asymptotische Verhalten des Nenners ist genauso wie bei dem Beweis der Konsistenz, also:

$$\operatorname*{plim}_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} (x_t - \mu)^2 = \sigma_x^2.$$

– Wegen des Zählers multipliziert man in (5.3) die Differenz $\hat{\beta}_1 - \beta_1$ mit $\sqrt{T}$ und erhält:

$$\sqrt{T}\left(\hat{\beta}_1 - \beta_1\right) = \frac{\sqrt{T}\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})u_t}{\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^2}$$
$$\xrightarrow{d} \frac{1}{\sigma_x^2}W \sim N\left(0, \frac{\sigma^2\sigma_x^2}{\left(\sigma_x^2\right)^2}\right) = N\left(0, \frac{\sigma^2}{\sigma_x^2}\right). \tag{5.5}$$

5.3.6. Wann ist ein Schätzer inkonsistent?

- **Einfaches Regressionsmodell**

  Im einfachen Regressionsmodell ist es ganz einfach: Wenn

  $$Cov(x, u) \neq 0.$$

  Dies passiert z.B., wenn TS.1 und damit TS.1' verletzt ist und das korrekte Modell lautet: $y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + v_t$. In diesem Fall sagt man, dass der OLS-Schätzer **asymptotisch verzerrt** ist.

- **Allgemeines Regressionsmodell**

  Für das allgemeine Regressionsmodell gilt, wenn $Cov(x_j, u) \neq 0$ für mindestens ein Regressor $j = 1, 2, \ldots, k$.

  Eine korrekte Modellwahl ist also für Konsistenz eine not-

wendige Voraussetzung! Genauer: die Wahl eines Modells, das alle relevanten Regressoren enthält, ist notwendige Voraussetzung für die Konsistenz eines Schätzers.

- **Messfehler**

  $Cov(x_j, u) = 0$ ist auch bei Messfehlern verletzt. Bei einer Variablen liegt ein **Messfehler** vor, wenn statt x_t lediglich z_t beobachtet werden kann und z.B.:

$$z_t = x_t + v_t, \quad v_t \sim \text{i.i.d.}(0, \sigma_v^2)$$

  gilt. In solchen Fällen ist der OLS-Schätzer nicht konsistent und es muss z.B. ein Instrumentvariablenschätzer angewendet werden, siehe die BA-Veranstaltung **Weiterführende Fragen der Ökonometrie** im Wintersemester.

5.4. Asymptotische Tests

5.4.1. t-Tests

Der Einfachheit halber wird zunächst das einfache Regressionsmodell für eine Zufallsstichprobe betrachtet:

$$y_t = \beta_0 + \beta_1 x_t + u_t, \quad t = 1, 2, \ldots, T.$$

Durch eine Standardisierung lässt sich (5.5) schreiben als:

$$\sqrt{T}\left(\frac{\hat{\beta}_1 - \beta_1}{\sigma/\sigma_x}\right) \xrightarrow{d} N(0,1).$$

Da im Allgemeinen σ und σ_x nicht bekannt sind, müssen sie geschätzt

werden. Hierfür sind die üblichen Schätzer geeignet:

$$\hat{\sigma} = \sqrt{\frac{1}{T-2}\sum_{t=1}^{T}\hat{u}_t^2}, \quad \hat{\sigma}_x = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(x_t - \bar{x})^2}.$$

Es lässt sich dann unter Verwendung obiger Resultate und der Rechenregeln relativ leicht zeigen, dass

$$\left(\sqrt{T}\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}\left(\hat{\sigma}_x\right)^{-1}}\right) \xrightarrow{d} N(0,1),$$

bzw.

$$\left(\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}\sqrt{\left(\sum_{t=1}^{T}(x_t - \bar{x})^2\right)^{-1}}}\right) \xrightarrow{d} N(0,1).$$

Der Term auf linken Seite der letzten Gleichung ist die t-**Statistik**!

Der t-Test gilt also nicht mehr exakt, sondern nur asymptotisch. Man spricht auch von einem **asymptotischen Test**.

Dies gilt unter den allgemeinen Bedingungen von Theorem 5.2, z.B. für streng stationäre AR(p)- oder dynamische Regressionsmodelle.

5.4.2. F-Tests

Es lässt sich zeigen, dass unter den Annahmen von Theorem 5.2 die F-Statistik asymptotisch $\mathrm{F}_{q,\infty} = \frac{1}{q}\chi^2(q)$ verteilt ist. Monte Carlo Simulationen haben gezeigt, dass die Approximation der (unbekannten) exakten Verteilung durch die $\mathrm{F}_{q,T-k-1}$-Verteilung besser ist.

Zu lesen: Sections 5.1, 5.2, 11.2, Appendix D und E.4 in Wooldridge (2009).

6. Nichtstationäre Zeitreihenprozesse: Random Walks und mehr

6.1. Vorhersagen (Teil I)

- Es gibt eine Reihe von Möglichkeiten, eine statistische Vorhersage (forecast) für einen zukünftigen Wert der abhängigen Variable, z.B. y_{t+h} auf Basis von Stichprobeninformation bis zum Zeitpunkt t zu machen. Es bezeichne I_t die **Informationsmenge** (information

set), die bis zum Zeitpunkt t für eine Vorhersage vorliegt. Im Fall einer Stichprobe einer univariaten Zeitreihe ist $I_t = \{y_1, y_2, \ldots, y_t\}$.

- Der **Vorhersagefehler** ist definiert als:

$$e_{t+h} = y_{t+h} - \text{Vorhersage von } y_{t+h} \text{ gegeben } I_t$$

- Bevor man eine Vorhersage machen kann, muss man die **Verlustfunktion** (loss function) wählen, mit der ein Vorhersagefehler bewertet wird. Beispiele sind:
  - Quadratischer Fehler: e^2_{t+h},
  - Absoluter Fehler: $|e_{t+h}|$.
- Die Verlustfunktion erlaubt die Bewertung des Vorhersagefehlers für *eine* Vorhersage. Um den Verlust einer Vorhersage "im Mittel" zu

bestimmen, betrachtet man den Erwartungswert des Verlustes. Für den quadratischen Verlust entspricht dies dem **Mittleren Quadratischen Fehler** (Mean Squared Error, MSE):

$$E[e_{t+h}^2|I_t] =$$
$$E\left[(y_{t+h} - \text{Prognose von } y_{t+h} \text{ gegeben Zeitreihenwerte } y_1, \ldots, y_t)^2 \,|I_t\right].$$

- Der MSE wird durch den bedingten Erwartungswert $E[y_{t+h}|I_t]$ minimiert. (Beweis ist nicht so schwer bzw. siehe Hinweise in Section 18.5 Wooldridge (2009).)

  Der bedingte Erwartungswert liefert also die optimale Vorhersage, wenn der MSE minimiert werden soll.

- Vorhersagen (Teil II); siehe Kapitel 11.

6.2. Random Walks: jetzt etwas genauer

Vgl. auch Diskussion in Abschnitt 4.2, insbesondere Abschnitt 4.2.6.

6.2.1. Begriff des Random Walks

Ein Random Walk ist ein AR(1)-Prozess mit $\alpha = 1$:

$$y_t = y_{t-1} + u_t, \quad u_t \sim WN(0, \sigma^2), \quad t = 1, 2, \ldots.$$

Alternative Darstellung als ungewichtete Summe eines White Noise Prozesses:

$$y_t = y_0 + \sum_{j=1}^{t} u_j. \tag{6.1}$$

Eine weitere nützliche Darstellung lautet (hier für allgemeine AR(1)-Prozesse):

$$\begin{aligned} y_{t+h} &= \alpha^h y_t + \alpha^{h-1} u_{t+1} + \cdots + \alpha u_{t+h-1} + u_{t+h} \\ &= \alpha^h y_t + \sum_{j=0}^{h-1} \alpha^j u_{t+h-j}. \end{aligned} \tag{6.2}$$

6.2.2. Eigenschaften

- **Erwartungswert**: $E[y_t] = E[y_0]$ und $E[y_t] = y_0$, falls y_0 nicht-stochastisch
- **Varianz**: $Var(y_t) = Var(y_0) + \sigma^2 t = \sigma^2 t$, falls $Var(y_0) = 0$.
- **Autokovarianz**: $Cov(y_t, y_{t-h}) = \sigma^2(t-h)$, $t > h \geq 0$.

- **Autokorrelation**: $h \geq 0$:
  $$Corr(y_{t+h}, y_t) = \frac{\sigma^2 t}{\sqrt{\sigma^2(t+h)\sigma^2 t}} = \sqrt{\frac{t}{t+h}}.$$
  - Für festes h gilt: $Corr(y_{t+h}, y_t) \overset{t\to\infty}{\longrightarrow} 1$.
  - Für jedes beliebige h findet sich ein t, so dass $Corr(y_{t+h}, y_t) > c$, $0 \leq c < 1$. Deshalb ist ein Random Walk nicht asymptotisch unkorreliert.
- Ein Random Walk ist ein **nichtstationärer Prozess**, da die Varianz und Kovarianzen sowie Korrelationen von der Zeit t abhängen.
- Ein Schock u_j bleibt in seiner Wirkung in alle Zukunft erhalten, denn gemäß (6.1) gilt:
  $$y_t = u_t + u_{t-1} + \cdots + u_j + \cdots + u_1 + y_0, \quad j < t.$$

Ein Random Walk ist deshalb ein Beispiel eines stochastischen Prozesses mit **starker Abhängigkeit** (strong dependence) bzw. **großer Persistenz** (high persistence) eines Schocks u_j, $j < t$ auf y_t. Dies zeigt sich auch darin, dass die Kovarianzen eines Random Walks für gegebenes t viel langsamer gegen Null (linear) gehen als die eines stationären AR(1)-Prozesses (exponentiell).

- Abbildung 6.1 zeigt zehn Realisationen von Random Walks, die mit dem R-Programm in Appendix A.12 erstellt wurden. Beim Betachten der Realisationen entsteht der Eindruck, als ob längere Abschnitte einen linearen Trend aufweisen würden. Meist ändert ein vermeintlicher linearer Trend nach einiger Zeit seine Richtung. Deshalb spricht man in diesem Fall auch von einem **stochastischen Trend**.

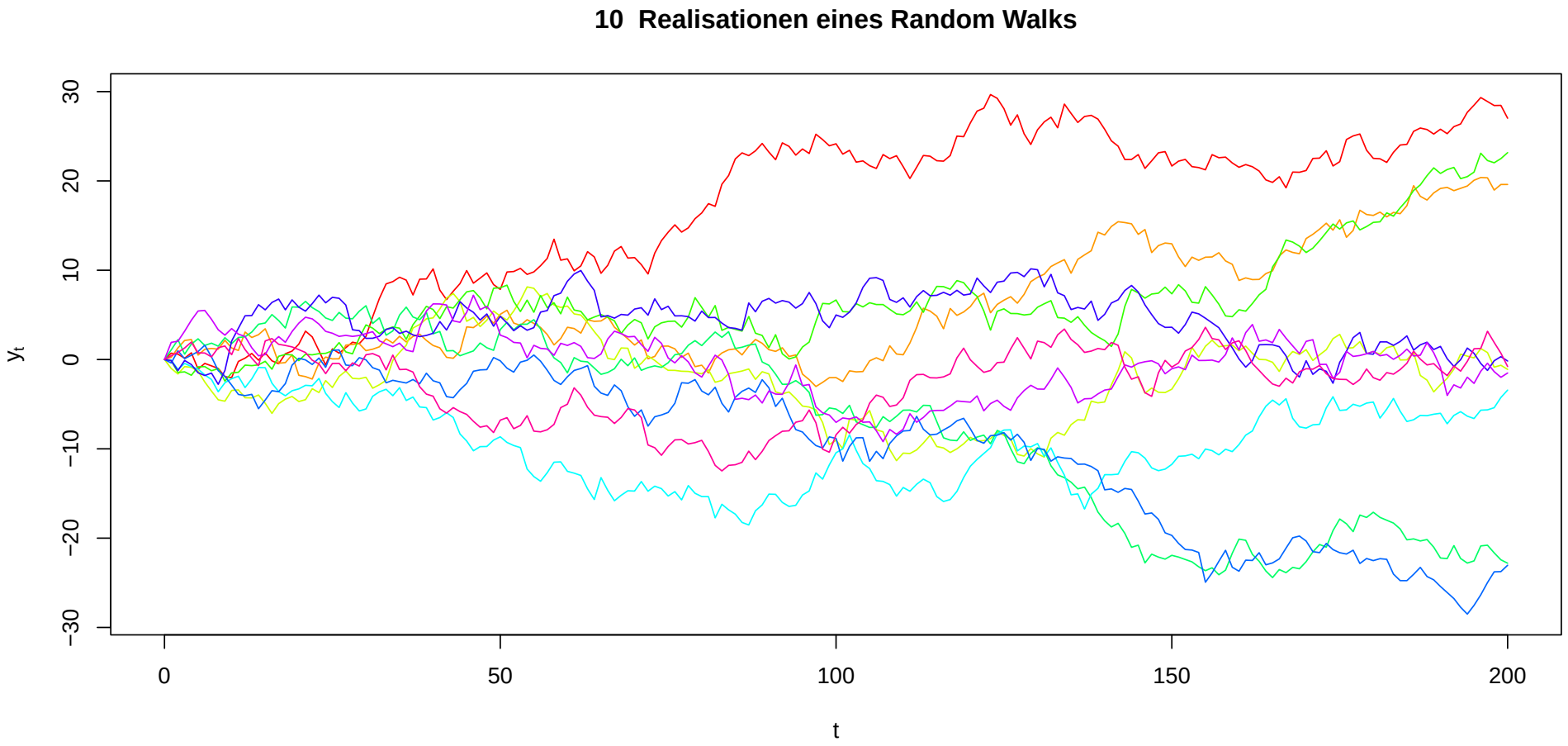

Abbildung 6.1.: Zehn Realisationen von Random Walks, siehe Appendix A.12 für das R-Programm.

6.2.3. Bedingte Momente und Vorhersagen

- Der **bedingte Erwartungswert** ergibt sich direkt aus (6.2) und den zusätzlichen Annahmen in Abschnitt 4.2.7:

$$E[y_{t+h}|y_t] = \alpha^h y_t \overset{h\to\infty}{\longrightarrow} \begin{cases} 0 & \text{falls AR(1) stationär: } |\alpha| < 1, \\ y_t & \text{falls AR(1) Random Walk: } \alpha = 1. \end{cases}$$

- Die **bedingte Varianz** (des Vorhersagefehlers) ergibt sich direkt aus (6.2):

$$Var(e_{t+h}|y_t) = Var(y_{t+h}|y_t) = \sigma^2 \sum_{j=0}^{h-1} \alpha^{2j}$$

$$\overset{h\to\infty}{\longrightarrow} \begin{cases} Var(y_t) = \frac{\sigma^2}{1-\alpha^2} & \text{falls AR(1) stationär,} \\ \lim_{h\to\infty} \sigma^2 h = \infty & \text{falls AR(1) RW.} \end{cases}$$

Beim stationären Prozess konvergiert diese hingegen gegen die unbedingte Varianz (vgl. Abschnitt 4.2.4).

- **Langfristige Vorhersagen** für stationäre und nichtstationäre AR-Prozesse unterscheiden sich grundlegend!
  - Bei stationären AR(1) Prozessen entspricht die langfristige Vorhersage (h groß) dem unbedingten Mittelwert und der Wert "heute", y_t, wird vergessen.
  - Beim Random Walk ist das Gegenteil der Fall. Der Wert "heute" ist die beste Vorhersage für jeden Vorhersagehorizont h.

  Und: Im Gegensatz zum stationären Prozess steigt beim Random Walk die Varianz des Vorhersagefehlers mit dem Vorhersagehorizont h unbegrenzt (linear) an.

6.2.4. Integrationsgrad stochastischer Prozesse

- Ein Random Walk ist eine ungewichtete Summe von Fehlern $\{u_t\}$, siehe (6.1). Da eine Summe ein Spezialfall eines Integrals ist, sagt man, dass ein Random Walk **integriert** ist **vom Grade 1 (integrated of order one)** oder **I(1)**. Der stochastische Prozess der Fehler ist hingegen **integriert vom Grade 0** bzw. **I(0)**.

- Es gilt auch, dass schwach abhängige stochastische Prozesse I(0) sind, z.B. stationäre AR(p)-Prozesse (ohne Beweis).

- Verallgemeinert man den RW, indem er statt durch i.i.d. Fehler durch einen schwach abhängigen Prozess $s_t \sim \mathsf{I}(0)$ "gefüttert" wird, gilt:

$$y_t = y_{t-1} + s_t \quad \Longrightarrow \quad y_t \sim \mathsf{I}(1).$$

Daraus folgt: die ersten Differenzen $y_t - y_{t-1} = s_t$ sind I(0).

- Ausblick: Es gibt auch Prozesse mit höheren Integrationsgraden, wobei $d \in \mathbb{N}$. Deren Darstellung ist am einfachsten möglich mit dem Differenzenoperator $\Delta y_t = y_t - y_{t-1}$. Für diesen gilt: $\Delta^2 = \Delta(\Delta y_t) = \Delta(y_t - y_{t-1}) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}$ und entsprechend für Δ^d. Ein Prozess y_t ist **integriert vom Grade** d

$$y_t \sim I(d),$$

wenn, gegeben $u_t \sim I(0)$, bzw. $s_t \sim I(0)$, gilt:

$$\Delta^d y_t = u_t$$
$$\Delta^d y_t = s_t$$

6.3. Random Walk mit Drift

Der Random Walk mit Drift ist eine Kombination eines stochastischen und eines deterministischen Trends:

$$\begin{aligned} y_t &= \mu_1 + y_{t-1} + u_t \\ &= \mu_1 t + \sum_{j=1}^{t} u_j + y_0. \end{aligned}$$

6.3.1. Eigenschaften Random Walk mit Drift

- Man beachte, dass die Konstante im Random Walk keinen konstanten Mittelwert impliziert wie im stationären AR-Prozess (vgl. hierzu

Abschnitt 4.2), sondern einen deterministischen linearen Trend:

$$E[y_t] = \mu_t = \mu_1 t + E[y_0].$$

Der Parameter μ_1 wird auch als **Drift** oder **Driftterm** und $\{y_t\}$ als **Random Walk mit Drift** bezeichnet: Je länger der Prozess läuft, desto weiter bzw. länger kann sich y_t vom linearen Trend entfernen. In `R` generieren!

- Bedingter Erwartungswert und Varianz des Vorhersagefehlers:

$$E[y_{t+h}|y_t] = y_t + \mu_1 h, \qquad \text{geg. Annahmen aus Absch. 4.2.7}$$
$$Var(y_{t+h}|y_t) = \sigma^2 h.$$

Liegen also ein stochastischer und ein deterministscher Trend vor, so nähert sich auch langfristig der bedingte Erwartungswert dem Zeittrend nicht an.
Darüber hinaus wächst die Vorhersagefehlervarianz unbegrenzt mit

h. Vgl. Vorhersage bei trendstationären AR(1)-Prozessen.

- Abbildung 6.2 zeigt zehn Realisationen eines Random Walks mit Drift, die mit dem R-Programm in Appendix A.12 erstellt wurden. In diesem Fall treten sowohl ein **deterministischer Trend** als auch ein **stochastischen Trend** auf.

- In der Praxis ist häufig nicht leicht zu unterscheiden, ob ein stochastischer Trend, ein deterministischer Trend oder beides vorliegt, siehe Kapitel 9.

6.3.2. Anmerkungen

- Diese Ergebnisse bleiben qualitativ erhalten, wenn statt e_t ein schwach abhängiger Prozess s_t vorliegt, siehe Kapitel 9.

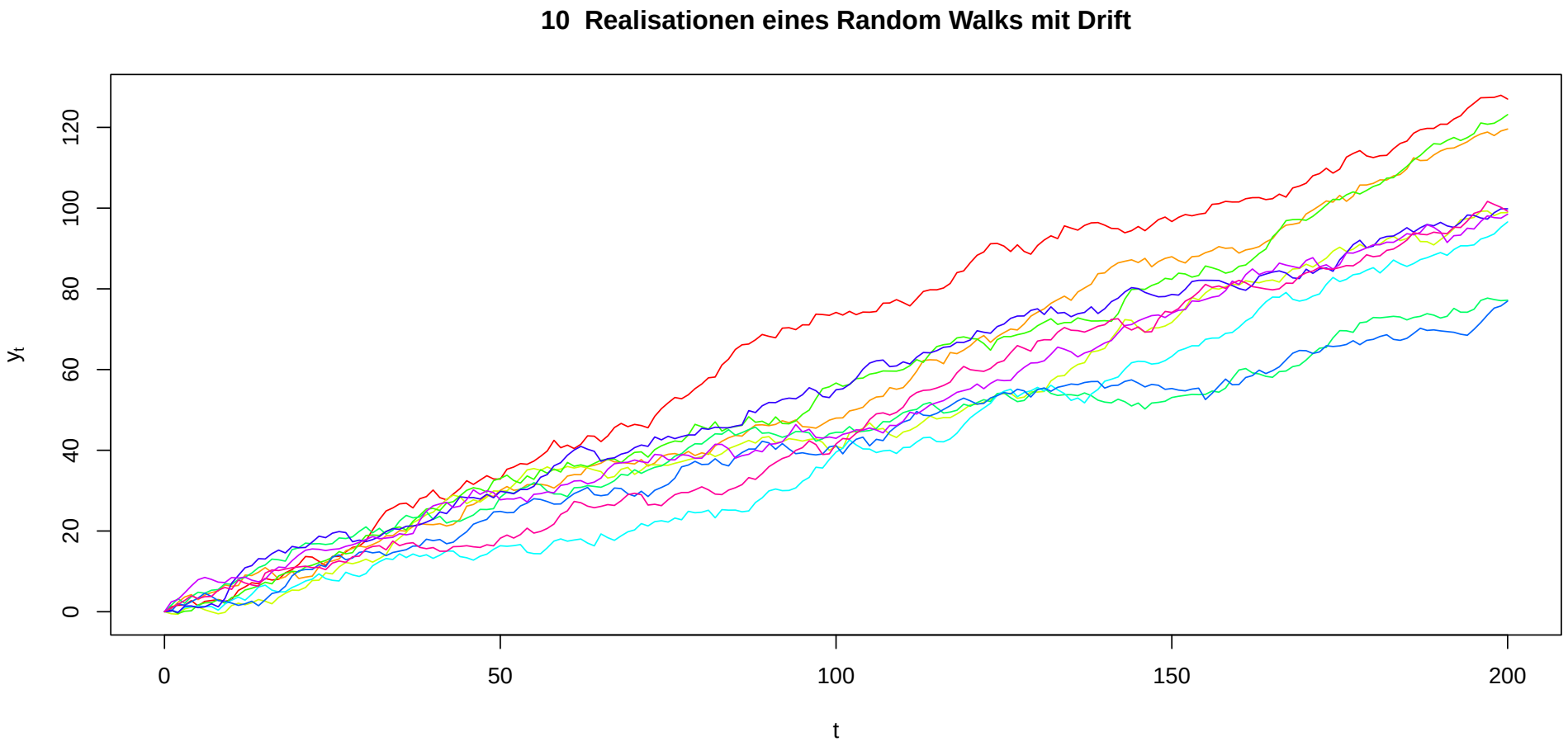

Abbildung 6.2.: Zehn Realisationen von Random Walks, siehe Appendix A.12 für das R-Programm.

- **Wachstumsraten** können durch Log-Differenzen approximiert werden:

$$\Delta \log(y_t) \approx \frac{y_t - y_{t-1}}{y_{t-1}}.$$

6.4. Regressionen mit I(1)-Variablen

Werden in einer Regression I(1)-Variablen verwendet, kann Annahme TS.1' verletzt sein. Dann gelten die asymptotischen Schätzeigenschaften des OLS-Schätzers aus Abschnitt 5.3 nicht oder z.T. nicht.

Einfacher Fall: $y_t \sim \mathsf{I}(1)$ und $x_t \sim \mathsf{I}(1)$:

$$y_t = \mu_1 + y_{t-1} + \tilde{y}_t, \quad \tilde{y}_t \sim \mathsf{I}(0) \text{ mit } E[\tilde{y}_t] = 0, \tag{6.3}$$

$$x_t = \delta_1 + x_{t-1} + \tilde{x}_t, \quad \tilde{x}_t \sim \mathsf{I}(0) \text{ mit } E[\tilde{x}_t] = 0. \tag{6.4}$$

In ersten Differenzen erhält man:

$$\Delta y_t = \mu_1 + \tilde{y}_t,$$
$$\Delta x_t = \delta_1 + \tilde{x}_t.$$

- Falls Δy_t und Δx_t als absolute Veränderungen interpretiert werden können, bezeichnen $\tilde{y}_t$ und $\tilde{x}_t$ die Abweichungen von den Erwartungswerten μ_1 und δ_1 der absoluten Veränderungen.

- Falls Δy_t und Δx_t als Wachstumsraten interpretiert werden können, bezeichnen $\tilde{y}_t$ und $\tilde{x}_t$ die Abweichungen von den erwarteten Wachstumsraten μ_1 und δ_1.

Elementare Fälle möglicher Regressionsbeziehungen

- Fall I: Linearer Zusammenhang mit stochastischem und evtl. auch deterministischem Trend,

- Fall II: Linearer Zusammenhang zwischen den Abweichungen,

- Fall III: $\beta_1 = 0$,

- Fall IV: Kointegrationsbeziehung mit Zeittrend.

6.4.1. Fall I Linearer Zusammenhang mit stochastischem und evtl. auch deterministischem Trend

Es gibt einen linearen Zusammenhang zwischen den Variablen mit stochastischen Trends und, falls $\mu_1 \neq 0$ oder $\delta_1 \neq 0$, ebenso mit deterministischen Trends:

$$y_t = \beta_0 + \beta_1 x_t + u_t, \quad u_t \sim WN(0, \sigma^2), \tag{6.5}$$

mit $\beta_1 \neq 0$.

Schätzeigenschaften des OLS-Schätzers für (6.5)

In diesem Fall weist der OLS-Schätzer für β_0 und β_1 (*außer* x_t *ist streng exogen*) eine **Nichtstandardverteilung** auf, d.h. er ist asym-

ptotisch nicht normalverteilt. Damit sind Standard t-Tests und F-Tests nicht anwendbar. Allerdings ist der OLS-Schätzer **konsistent**. Außerdem führt die **Verletzung von Annahme TS.3'** nicht unbedingt zu einem inkonsistenten Schätzer. Siehe Kapitel 10 für Details.

Alternative Schätzgleichung

Damit der OLS-Schätzer für β_1 asymptotisch normalverteilt ist, ist eine notwendige (jedoch nicht hinreichende) Bedingung, dass (6.5) umgeformt wird, indem die Gleichungen (6.3) und (6.4) in (6.5) eingesetzt werden:

$$\mu_1 + y_{t-1} + \tilde{y}_t = \beta_0 + \beta_1 (\delta_1 + x_{t-1} + \tilde{x}_t) + u_t.$$

Man kann dies umformen zu

$$\mu_1 + \tilde{y}_t = \beta_0 + \underbrace{(\beta_1 x_{t-1} - y_{t-1})}_{=-(\beta_0+u_{t-1})\sim \mathsf{I}(0)} + \beta_1(\delta_1 + \tilde{x}_t) + u_t,$$

bzw. zu

$$\Delta y_t = \underbrace{\beta_0 + (\beta_1 x_{t-1} - y_{t-1})}_{=-u_{t-1}\sim \mathsf{I}(0)} + \beta_1 \Delta x_t + u_t. \tag{6.6}$$

- Man beachte die Besonderheit des Terms:

  $$(\beta_1 x_{t-1} - y_{t-1}) = -(\beta_0 + u_{t-1}) \sim \mathsf{I}(0).$$

  Diese Linearkombination zweier I(1)-Variablen ist I(0), da $u_t \sim WN(0, \sigma^2) \sim \mathsf{I}(0)$ und damit schwach abhängig ist.

- Linearkombinationen zwischen I(1)-Variablen, die selbst I(0) sind, werden als **Kointegrationsbeziehungen** bezeichnet. Kointegrationsbeziehungen spiegeln in der Ökonomie häufig **langfristige**

Gleichgewichtsbeziehungen wider und spielen deshalb in der modernen Zeitreihenanalyse eine essentielle Rolle!

- Schreibt man die Linearkombination $(\beta_1 x_{t-1} - y_{t-1})$ in Vektorform:

$$\begin{pmatrix} \beta_1 & -1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix},$$

wird der Vektor $\begin{pmatrix} \beta_1 & -1 \end{pmatrix}$ als **Kointegrationsvektor** bezeichnet.

Man beachte, dass der Kointegrationsvektor **nicht eindeutig** ist, denn

$$-\begin{pmatrix} -\beta_1 & 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} = -\alpha \begin{pmatrix} \frac{-\beta_1}{\alpha} & \frac{1}{\alpha} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix}$$

spiegeln die gleiche Kointegrationsbeziehung wider.

- Eine Regression von Δy_t auf Δx_t ist im Fall I unvollständig spezifiziert, wenn als Regressor die **Abweichung vom langfristigen Gleichgewicht** fehlt! (Dies könnte zum Beispiel auch im Beispiel 11.6 in Wooldridge (2009) der Fall sein.)

- Der Parameter α gibt hier die **Anpassungsgeschwindigkeit** an das langfristige Gleichgewicht an (multivariate Entsprechung: **loading matrix**). Um ein eindeutiges α zu erhalten, normiert man α häufig so, dass y_{t-1} mit Faktor Eins in die Kointegrationsbeziehung eingeht. Man erhält aus (6.6)

$$\Delta y_t = \beta_0 + (-1)\,(y_{t-1} - \beta_1 x_{t-1}) + \beta_1 \Delta x_t + u_t$$

und somit $\alpha = -1$. Ist der DGP unbekannt, schreibt man allgemein

$$\Delta y_t = \beta_0 + \alpha\,(y_{t-1} - \beta_1 x_{t-1}) + \beta_1 \Delta x_t + u_t. \qquad (6.7)$$

Da hierdurch Abweichungen vom langfristigen Gleichgewicht (d.h.

"Fehler") korrigiert werden und dies zu einer Beeinflussung der Änderung alias Differenzen Δy_t führt, wird die Darstellung (6.7) als **Fehlerkorrektur-Modell** bezeichnet.

- **OLS-Schätzung des Fehlerkorrektur-Modells (6.7)**:
  - Ist der Kointegrationsvektor **bekannt**, sind die abhängige Variable Δy_t und alle Regressoren, also $\beta_1 x_{t-1} - y_{t-1}$ und Δx_t I(0)-Variablen und die Annahme TS.1' ist erfüllt.

    Achtung: Damit der OLS-Schätzer konsistent ist, müssen auch die Annahmen TS.2' und TS.3' erfüllt sein. TS.3' ist jedoch nicht erfüllt, wenn z.B. $E[u_t|\Delta x_t] \neq 0$ ist. Dies ist z.B. der Fall, wenn u_t und Δx_t korreliert sind. Für die asymptotische Normalverteilung müssen auch noch TS.4' und TS.5' gelten.

– Ist der Kointegrationsvektor **unbekannt**, kann dieser auf Basis von (6.6) geschätzt werden, indem man

$$\Delta y_t = \gamma_0 + \gamma_1 x_{t-1} + \gamma_2 y_{t-1} + \gamma_3 \Delta x_t + u_t \qquad (6.8)$$

schätzt. Dann ist der OLS-Schätzer für γ_3 asymptotisch normalverteilt, da Δx_t eine I(0)-Variable ist. Die OLS-Schätzer für γ_0, γ_1 und γ_2 haben keine Standardverteilung. Siehe Kapitel 10.

Aber Achtung: Unabhängig davon, ob die Kointegrationsbeziehung geschätzt werden muss, müssen auch die Annahmen TS.2' und TS.3' erfüllt sein, damit der OLS-Schätzer für γ_3 konsistent ist. Für die asymptotische Normalverteilung müssen auch noch TS.4' und TS.5' gelten. Weiter wie oben.

Jedoch bleiben die OLS-Schätzer für γ_0, γ_1 und γ_2 konsistent (ohne Nachweis).

6.4.2. Fall II: Linearer Zusammenhang zwischen den Abweichungen

Es gibt einen linearen Zusammenhang zwischen den Abweichungen:

$$\tilde{y}_t = \beta_0 + \beta_1 \tilde{x}_t + u_t, \quad u_t \sim WN(0, \sigma^2), \quad \beta_1 \neq 0.$$

Damit ergibt sich:

$$\begin{aligned} \Delta y_t - \mu_1 &= \beta_0 + \beta_1(\Delta x_t - \delta_1) + u_t, \\ \Delta y_t &= \beta_0 + \mu_1 - \beta_1\delta_1 + \beta_1 \Delta x_t + u_t = \nu + \beta_1 \Delta x_t + u_t \end{aligned}$$

und eine Regression von Δy_t auf Δx_t ist korrekt spezifiziert.

Schätzung: Man beachte wieder, dass auch TS.2' und TS.3' für die Konsistenz des OLS-Schätzers erfüllt sein müssen (Gründe für Verletzung siehe oben). Für die asymptotische Normalverteilung müssen auch noch TS.4' und TS.5' gelten.

6.4.3. Fall III: $\beta_1 = 0$

In Fall I und Fall II ist $\beta_1 \neq 0$.

- Dann ist eine Regression von Δy_t auf Δx_t unproblematisch - sofern auch TS.2' bis TS.5' erfüllt sind -, da man eine Regression zwischen I(0)-Variablen durchführt.

- Eine Regression von y_t auf x_t ist im Allgemeinen völlig irreführend, da in diesem Fall der OLS-Schätzer inkonsistent ist, da auch in sehr großen Stichproben die Varianz des Schätzers nicht verschwindet. Es liegt der Fall einer **Scheinregression** vor. Man kann also auch in großen Stichproben einen Zusammenhang zwischen y_t und x_t erhalten, obwohl es keinen Zusammenhang gibt.

Im Gegensatz zum Fall III bei der Regression mit deterministischen Trends (Abschnitt 3.3) kann dieses Problem nicht durch Berücksichtigung eines Zeittrends behoben werden!

Zusammenfassung: In diesem einfachen Spezialfall ist die Schätzung von (6.8) robust: γ_1 und γ_2 werden konsistent geschätzt und γ_3 ist asymptotisch normalverteilt, wenn TS.1' bis TS.3' gelten.

Beispiel: Keynesianische Konsumfunktion

$x_t = \log(Y_t)$ ist das logarithmierte Volkseinkommen, $y_t = \log(C_t)$ ist der logarithmierte aggregierte Konsum. Dann stellt:

$$\log(C_t) = \beta_0 + \beta_1 \log(Y_t) + u_t$$

eine **langfristige** Konsumfunktion dar! Die **Wachstumsraten des Konsums** $\Delta \log(C_t)$ hängen damit nicht nur von der Wachstumsrate des Einkommens, sondern auch von der Abweichung des Konsums vom langfristigen Gleichgewicht ab. Liegt das einfache Modell (6.5) zugrunde, könnte man β_1 durch (6.7) schätzen:

$$\Delta \log(C_t) = \beta_0 + \alpha(\log(C_{t-1}) - \beta_1 \log(Y_{t-1})) + \beta_1 \Delta \log(Y_t) + u_t.$$

6.4.4. Fall IV: Kointegrationsbeziehung mit Zeittrend

Gleichungen (6.3), (6.4) und (6.5) schließen einen Zeittrend in der Kointegrationsbeziehung implizit aus. Um dies zu erkennen, schreibt man (6.3) und (6.4) um, so dass der Zeittrend und der stochastische Trend getrennt werden:

$$y_t = \mu_1 t + \sum_{j=1}^{t} \tilde{y}_j,$$

$$x_t = \delta_1 t + \sum_{j=1}^{t} \tilde{x}_j.$$

O.E.d.A. wurden $y_0 = 0$ und $x_0 = 0$ gesetzt.

Da gemäß (6.5) $y_t - \beta_0 - \beta_1 x_t = u_t \sim WN(0, \sigma^2)$, muss dies auch

gelten, wenn y_t und x_t eingesetzt werden. Man erhält:

$$\begin{aligned} y_t - \beta_0 - \beta_1 x_t &= \mu_1 t + \sum_{j=1}^{t} \tilde{y}_j - \beta_0 - \beta_1 \left(\delta_1 t + \sum_{j=1}^{t} \tilde{x}_j \right) \\ &= -\beta_0 + \underbrace{(\mu_1 - \beta_1 \delta_1)}_{\text{A}} t + \underbrace{\sum_{j=1}^{t} \tilde{y}_j - \beta_1 \sum_{j=1}^{t} \tilde{x}_j}_{\text{B}} = u_t. \end{aligned}$$

Damit dass letzte Gleichheitszeichen gilt, muss der Ausdruck A gerade Null sein, da u_t sonst über $\text{A} \cdot t$ einen Trend enthält. Mit anderen Worten: der Zeittrend μ_1 in y_t wird gerade durch den Zeittrend δ_1 in x_t mal den Regressionsparameter β_1 bestimmt. In diesem Fall eliminiert der Kointegrationsvektor auch die jeweiligen Drifts bzw. linearen Trends, die x_t und y_t aufweisen, aus der Kointegrationsbeziehung.

Im "Zeitreihendialekt" sagt man auch, dass der *Trend orthogonal zur*

Kointegrationsbeziehung ist. Dies ist ganz analog zum Fall I in Abschnitt 3.3.

Da die linke Seite I(0) ist, muss B ebenfalls I(0) sein. Damit spezifiziert B wieder eine Kointegrationsbeziehung.

In der Praxis ist dieser Spezialfall nicht notwendigerweise gegeben. Dann muss (6.5) durch

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 t + u_t \tag{6.9}$$

ersetzt werden. Nun entspricht nach entsprechender Umformung $\beta_2 = \mu_1 - \beta_1 \delta_1$. Die OLS-Schätzer der Parameter β_i, $i = 0, 1, 2$ weisen wiederum eine Nichtstandardverteilung auf.

6.4.5. Anmerkungen

- Besteht eine Kointegrationsbeziehung der Form (6.5) oder (6.9) jedoch mit $u_t \sim$ I(0) statt $u_t \sim WN(0, \sigma^2)$, dann bleibt der OLS-Schätzer für die *Regression in Niveaus konsistent*, jedoch ändert sich wieder die spezifische Form der Nichtstandardverteilung.

- Für die *konsistente OLS-Schätzung in ersten Differenzen* ist die Annahme TS.3':

$$E[u_t|\mathbf{x}_t] = 0$$

zentral. Wenn diese Annahme verletzt ist, ist der OLS-Schätzer inkonsistent. Wann ist diese Annahme verletzt?

  – *Endogenität der Regressoren*: Beispiel:

$$x_t = \alpha x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2), \quad Cov(u_t, \varepsilon_t) \neq 0.$$

Können Sie sich ökonomische Beispiele vorstellen, wo dies der Fall ist? Adäquate Schätzmethoden, z.B. der Instrumentvariablenschätzer (IV-Schätzer) werden in **Weiterführende Fragen der Ökonometrie** vorgestellt.

– Gewisse Formen von *dynamisch unvollständiger Spezifikation*, siehe folgenden Kapitel 7.

- Für die *Schätzung der Schätzvarianz* sind auch die Annahmen TS.4' (Homoskedastidizität) und die Annahme TS.5' (keine Autokorrelation in den Fehlern) notwendig. TS.5' ist erfüllt, wenn das Modell dynamisch korrekt spezifiziert ist, siehe folgenden Kapitel 7. Alternativ lassen sich entsprechend robuste Schätzer verwenden, siehe Kapitel 8.

Zu Lesen: Section 11.3, Section 18.3 und 18.4 (ohne Schätzer und

Tests) in Wooldridge (2009).

Anhang zu Kapitel 6

Die Bildung erster Differenzen **beseitigt auch einen deterministischen linearen Trend** und erscheint damit als eine Alternative zur direkten Modellierung deterministischer Trends wie im Abschnitt 3.1:

$$y_t = \mu_0 + \mu_1 t + e_t, \quad e_t \sim \text{i.i.d.}(0, \sigma^2)$$
$$\Delta y_t = \mu_1 + e_t - e_{t-1} = \mu_1 + \underbrace{\Delta e_t}_{u_t} = \mu_1 + u_t.$$

Aber:

$$\begin{aligned} Cov(u_t, u_{t-1}) &= Cov(e_t - e_{t-1}, e_{t-1} - e_{t-2}) \\ &= Cov(-e_{t-1}, e_{t-1}) = -\sigma^2. \end{aligned}$$

Damit ist TS.5' (Abschnitt 5.3) verletzt! Der Fehlerprozess $\{u_t\}$ der ersten Differenzen ist autokorreliert und entspricht einem MA(1)-Prozess (vgl. Abschnitt 4). Für Lösungen siehe Abschnitt 8. Dieses Problem tritt nicht auf, wenn y_t ein Random Walk (mit Drift) ist.

7. Regressionsmodelle für Zeitreihendaten III: Dynamische Regressionsmodelle

7.1. Dynamisches Regressionsmodell

Für einen exogenen Regressor z_t und die endogene Variable y_t ist ein dynamisches Regressionsmodell gegeben durch:

$$\begin{aligned} y_t = \nu &+ \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} \\ &+ \delta_0 z_t + \delta_1 z_{t-1} + \cdots + \delta_q z_{t-q} + u_t, \end{aligned}$$

wobei natürlich noch weitere Regressoren kontemporär oder verzögert in das Modell eingehen können. Mit

$$\mathbf{x}_t = \begin{pmatrix} 1 & y_{t-1} & \cdots & y_{t-p} & z_t & z_{t-1} & \cdots & z_{t-q} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \nu \\ \alpha_1 \\ \vdots \\ \alpha_p \\ \delta_0 \\ \delta_1 \\ \vdots \\ \delta_q \end{pmatrix}$$

ergibt sich wieder das Regressionsmodell in Matrixschreibweise:

$$y_t = \mathbf{x}_t \boldsymbol{\beta} + u_t.$$

7.2. WDH: Verzerrung durch vergessene Regressoren

Wiederholung aus **Einführung in die Ökonometrie**, Abschnitt 3.3: Omitted Variable Bias.
Wird statt des korrekt spezifizierten Modells

$$\mathbf{y} = \mathbf{X}_A \boldsymbol{\beta}_A + \mathbf{x}_a \beta_a + \mathbf{u}$$

das Modell

$$\mathbf{y} = \mathbf{X}_A \boldsymbol{\beta}_A + \mathbf{w}$$

geschätzt, dann gilt:

$$\tilde{\boldsymbol{\beta}}_A = \boldsymbol{\beta}_A + \underbrace{(\mathbf{X}'_A \mathbf{X}_A)^{-1} \mathbf{X}'_A \mathbf{x}_a}_{\tilde{\boldsymbol{\delta}}} \beta_a + (\mathbf{X}'_A \mathbf{X}_A)^{-1} \mathbf{X}'_A \mathbf{u},$$

wobei $\tilde{\boldsymbol{\delta}}$ der KQ-Schätzer für die Regression von $\mathbf{x}_a$ auf $\mathbf{X}_A$ ist:

$$\mathbf{x}_a = \mathbf{X}_A \boldsymbol{\delta} + \boldsymbol{\varepsilon}.$$

Damit $\tilde{\boldsymbol{\beta}}_A$ konsistent sein kann, muss $\boldsymbol{\delta} = 0$ oder $\beta_a = 0$ sein. $\boldsymbol{\delta} = 0$ bedeutet, dass $\mathbf{x}_a$ und $\mathbf{X}_A$ unkorreliert sind. Letzeres ist auch Voraussetzung dafür, dass $E[\mathbf{w}|\mathbf{X}_A] = 0$. Im vorliegenden Fall ist jedoch:

$$\begin{aligned} E[\mathbf{w}|\mathbf{X}_A] &= E[\mathbf{x}_a\beta_a + \mathbf{u}|\mathbf{X}_A] = E[\mathbf{x}_a\beta_a|\mathbf{X}_A] + E[\mathbf{u}|\mathbf{X}_A] \\ &= E[\mathbf{x}_a|\mathbf{X}_A]\beta_a + \underbrace{E[\mathbf{u}|\mathbf{X}_A]}_{=0, \text{ da } E[\mathbf{u}|\mathbf{X}_A,\mathbf{x}_a]=0} = E[\mathbf{x}_a|\mathbf{X}_A]\beta_a \end{aligned}$$

und für $Cov(\mathbf{x}_a, \mathbf{X}_A) \neq 0$ folgt, dass $E[\mathbf{x}_a|\mathbf{X}_A] \neq 0$, so dass dann TS.3 verletzt ist.

Zur Berechnung von $E[\mathbf{u}|\mathbf{X}_A]$:

$$E[\mathbf{u}|\mathbf{X}_A] = E\left[E[\mathbf{u}|\mathbf{X}_A, \mathbf{x}_a]|\mathbf{X}_A\right] = E[0|\mathbf{X}_A] = 0,$$

wegen $E[\mathbf{u}|\mathbf{X}_A, \mathbf{x}_a] = 0$ gemäß Annahme.

Man beachte, dass:

$$E[\mathbf{w}|\mathbf{X}_A] = 0 \quad \Leftrightarrow \quad E[\mathbf{y}|\mathbf{X}_A] = \mathbf{X}_A\boldsymbol{\beta}_A.$$

7.3. Verzerrung durch vergessene Regressoren bei dynamischen Regressionsmodellen

Bei Regressionsmodellen für Zeitreihendaten kann es noch komplizierter werden.

7.3.1. Einfachstes Beispiel

Ein **statisches Regressionsmodell mit streng exogenem Regressor und autokorrelierten Fehlern**

$$y_t = \beta_0 + \beta_1 x_t + u_t, \quad u_t | x_1, \ldots, x_T \sim (0, \sigma^2),$$
$$u_t = \rho u_{t-1} + v_t, \quad v_t \sim \text{i.i.d.}(0, \sigma_v^2),$$
$$v_t \text{ und } x_s,\ t, s = 1, 2, \ldots, \text{ sind stochastisch unabhängig.}$$

Setzt man in den AR(1)-Prozess die Fehler $u_j = y_j - \beta_0 - \beta_1 x_j$, $j = t, t-1$ ein und formt um, erhält man:

$$y_t = \beta_0(1-\rho) + \rho y_{t-1} + \beta_1 x_t - \beta_1 \rho x_{t-1} + v_t.$$

Offensichtlich ist für das statische Regressionsmodell die Annahme TS.3 erfüllt, da $E[y_t|x_t] = \beta_0 + \beta_1 x_t$. Dagegen gilt:

$$\begin{aligned} E[y_t|\mathbf{x}_t, y_{t-1}, \mathbf{x}_{t-1}, y_{t-2}, \ldots] &= \beta_0(1-\rho) + \rho y_{t-1} + \beta_1 x_t - \beta_1 \rho x_{t-1} \\ &\neq \beta_0 + \beta_1 x_t = E[y_t|x_t]. \end{aligned}$$

Obwohl das statische Modell korrekt spezifiziert ist, da $E[u_t|x_t] = 0$, ist das statische Modell aus dynamischer Sicht fehlspezifiziert, da sich der bedingte Erwartungswert unterscheidet, wenn zusätzliche Lags in das Modell aufgenommen werden.

7.3.2. Dynamisch vollständig spezifiziertes Modell

Wenn dagegen die Erwartungswerte übereinstimmen, wird ein Regressionsmodell als **dynamisch vollständig spezifiziertes Modell** (dynamically complete model) bezeichnet, d.h.:

$$E[y_t|\mathbf{x}_t, y_{t-1}, \mathbf{x}_{t-1}, y_{t-2}, \ldots] = E[y_t|\mathbf{x}_t].$$

Ist ein Modell **dynamisch vollständig spezifiziert, ist automatisch Annahme TS.5' erfüllt** (Beweis siehe Übung).

- **Bei Zeitreihendaten kann das Vergessen von (verzögerten) Regressoren zweierlei bewirken**:

  A) Autokorrelation in den Fehlern (neu),

  B) inkonsistente Schätzung der gewünschten Parameter.

7.3.3. Einfaches Beispiel für Fall B

Die Stichprobe wird von einem **stationären AR(2)-Prozess** generiert:

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + u_t, \quad u_t \sim WN(0, \sigma^2),$$

mit $\alpha_1, \alpha_2 \neq 0$, so dass

$$E[y_t | y_{t-1}, y_{t-2}, y_{t-3}, \ldots] = \alpha_1 y_{t-1} + \alpha_2 y_{t-2},$$

also

$$E[y_t | y_{t-1}, y_{t-2}] = \alpha_1 y_{t-1} + \alpha_2 y_{t-2}$$

das dynamisch vollständig spezifizierte Modell ist.

Schätzt man stattdessen ein AR(1)-Modell:

$$y_t = \alpha_1 y_{t-1} + w_t,$$

dann ist

$$w_t = \alpha_2 y_{t-2} + u_t$$

und

$$E[w_t|y_{t-1}] = E[\alpha_2 y_{t-2} + u_t|y_{t-1}] = \alpha_2 E[y_{t-2}|y_{t-1}] + \underbrace{E[u_t|y_{t-1}]}_{=0}$$

$$= \alpha_2 E[y_{t-2}|y_{t-1}] \neq 0,$$

da bei einem AR(2)-Prozess i.A. $Cov(y_{t-1}, y_{t-2}) \neq 0$ ist.

$\Rightarrow$ Damit ist TS.3' verletzt und α_1 des AR(2)-Prozesses kann mit einem AR(1)-Modell nicht konsistent geschätzt werden!

Aber: Man kann die Sache auch anders sehen. Betrachtet man

$$y_t = a y_{t-1} + v_t$$

und berücksichtigt, dass $E[y_{t-2}|y_{t-1}] = \rho y_{t-1}$ (ohne Beweis), dann ergibt sich $a = \alpha_1 + \alpha_2 \rho$. Der Parameter a wird konsistent geschätzt

und TS.3' ist hinsichtlich a erfüllt. Gleichzeitig weisen die Fehler v_t Autokorrelation erster Ordnung auf.

7.3.4. Einfaches Beispiel für Fall B bei einem dynamischen Regressionsmodell

Statt des dynamisch vollständig spezifizierten Modells

$$y_t = \alpha_1 y_{t-1} + \delta_1 z_t + u_t$$

wird das Modell

$$y_t = \beta z_t + v_t$$

geschätzt. Dann wird der Parameter zu z_t inkonsistent geschätzt, wenn $E[y_{t-1}|z_t] \neq 0$. Es stellt sich also wiederum die Frage, ob man an δ_1 oder $\beta \neq \delta_1$ interessiert ist.

Fazit

Um die Schätzeigenschaften beurteilen zu können, kommt es also darauf an, welcher Parameter – hier α_1 oder a, bzw. δ_1 oder β – für den Modellierungszweck relevant ist. Ist demnach die Annahme TS.3' erfüllt (Parameter a bzw. β sind relevant), dann bleibt die Verletzung der Annahme TS.5' durch die dynamisch unvollständige Spezifikation. Möglichkeiten zur Schätzung werden im Kapitel 8 behandelt.

Allerdings sollte man, sofern die Stichprobengröße es erlaubt, möglichst versuchen, ein dynamisch vollständig modelliertes Modell zu erstellen.

Anmerkung: Die Annahme TS.4' **homoskedastischer Fehler** kann auch verletzt sein. Dies wird ebenfalls im Kapitel 8 behandelt.

Zu lesen: Sections 11.4, 12.1, Seite 411f in Wooldridge (2009).

8. Regressionsmodelle für Zeitreihendaten IV: Regressionsmodelle mit autokorrelierten und heteroskedastischen Fehlern

Autokorrelierte Fehler in Regressionsmodellen

- Die exakte Wirkung autokorrelierter Fehler in Regressionsmodellen hängt davon ab, ob die Regressoren streng exogen sind oder nicht.
- In beiden Fällen dürfen die bekannten Standardfehler für die Parameterschätzer nicht verwendet werden, also auch nicht zur Berechnung von t-Statistiken!

8.1. Streng exogene Regressoren

8.1.1. Modell

$$y_t = \beta_1 x_t + u_t,$$
$$u_t = \rho u_{t-1} + e_t, \quad e_t \sim \text{i.i.d.}(0, \sigma_e^2).$$

Die Konstante β_0 wurde zur Vereinfachung des KQ-Schätzers weggelassen. (Sonst müsste im Folgenden x_t durch $x_t - \bar{x}$ ersetzt werden.)

8.1.2. Schätzung

Dann ist

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{t=1}^{T} x_t u_t}{\sum_{t=1}^{T} x_t^2}.$$

8.1.3. Verzerrte Standardfehler

Die Varianz der Parameterschätzung lautet, wegen $Cov(u_t, u_{t+j}|\mathbf{X}) = \rho^j \sigma^2$:

$$Var(\hat{\beta}_1|\mathbf{X}) = \frac{1}{\left(\sum_{t=1}^T x_t^2\right)^2} Var\left(\sum_{t=1}^T x_t u_t|\mathbf{X}\right)$$

$$= \frac{1}{\left(\sum_{t=1}^T x_t^2\right)^2} \left(\sum_{t=1}^T x_t^2 Var(u_t|\mathbf{X}) + 2\sum_{t=1}^{T-1}\sum_{j=1}^{T-t} x_t x_{t+j} Cov(u_t, u_{t+j}|\mathbf{X})\right)$$

$$= \frac{\sigma^2}{\left(\sum_{t=1}^T x_t^2\right)} + \underbrace{2\frac{\sigma^2}{\left(\sum_{t=1}^T x_t^2\right)^2}\sum_{t=1}^{T-1}\sum_{j=1}^{T-t} \rho^j x_t x_{t+j}}_{\text{zusätzlich bei AR-Fehlern}}.$$

Falls die unabhängige Variable x_t positiv autokorreliert ist, gilt:

- $\rho > 0$: Standardfehler wird bei gewöhnlichem KQ-Schätzer **unterschätzt**.
- $\rho < 0$: Standardfehler kann bei gewöhnlichem KQ-Schätzer **überschätzt** werden.

Man beachte, dass in beiden Fällen das geschätzte R^2 wie bisher interpretiert werden kann, da R^2 die Fehlervarianz mit der Varianz von y_t vergleicht und beide Varianzen im Falle stationärer, schwach abhängiger Prozesse konsistent geschätzt werden.

Zu lesen: Chapter 12.1, Seiten 408-410 in Wooldridge (2009).

8.1.4. Tests auf Autokorrelation erster Ordnung

Aufgrund der strengen Exogenität der Regressoren gilt $E[\mathbf{u}|\mathbf{X}] = 0$ und weiter gilt aufgrund der angenommen Modellstruktur:

$$E[e_t|u_{t-1}, u_{t-2}, \ldots] = 0, \quad Var(e_t|u_{t-1}, \ldots) = \sigma_e^2.$$

- **Asymptotischer t-Test**
  Liegt im genannten Beispiel keine Autokorrelation vor, ist $\rho = 0$. Man testet deshalb in

$$u_t = \rho u_{t-1} + e_t, \quad t = 2, \ldots, T,$$

$$\mathsf{H}_0 : \rho = 0 \text{ gegen } \mathsf{H}_1 : \rho \neq 0,$$

  wobei $|\rho| < 1$ unter H_1 vorausgesetzt werden muss. Warum?

  – Da der Fehlerprozess $\{u_t\}$ unbekannt ist, wird er durch den Pro-

zess der Residuen der Regression von y_t auf eine Konstante und x_t geschätzt. Die Testregression lautet dann:

$$\hat{u}_t = \rho \hat{u}_{t-1} + \varepsilon_t.$$

Es kann gezeigt werden (schwierig), dass trotz der Verwendung der Residuen gilt:

$$t_\rho \xrightarrow{d} N(0,1).$$

Die Intuition für dieses Ergebnis ist, dass $\text{plim}_{T\to\infty} \hat{u}_t = u_t$ gilt, wenn β_1 konsistent geschätzt wird.

- Dieses Ergebnis bleibt erhalten, wenn x_t durch einen Vektor $\mathbf{x}_t$ streng exogener Regressoren ersetzt wird.

- Der Test hat (asymptotisch) Güte (Macht, Power) gegen jede Form der Autokorrelation in den Fehlern, sofern $Cov(u_t, u_{t-1}) \neq 0$.

- Geht man davon aus, dass unter H_1 die Autokorrelation positiv ist, sollte man einen einseitigen Test durchführen, um die Power zu erhöhen.

- In Appendix A.14 findet sich R-Programm, das die Durchführung des Tests anhand von simulierten Stichprobendaten illustriert.

- **Durbin-Watson Test**

Die Durbin-Watson Teststatistik lautet:

$$DW = \frac{\sum_{t=2}^{T} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{T} \hat{u}_t^2}.$$

– Folgen die Fehler einem AR(1)-Prozess, gilt:

$$DW \approx 2(1-\rho) \quad \text{bzw.} \quad \rho \approx 1 - \frac{DW}{2},$$

so dass allein deshalb die Angabe der DW-Statistik im Software-Output hilfreich ist.

– Die Hypothese für den DW-Test ist standardmäßig einseitig $H_0 : \rho \leq 0$ versus $H_1 : \rho > 0$, kann jedoch auch für den gegenteiligen Fall formuliert werden.

– Die Verteilung der DW-Statistik hängt von $\mathbf{X}$ ab. Es kann jedoch eine obere und untere Schranke angegeben werden (siehe Section 12.2, Seiten 415-416 in Wooldridge (2009)).

Zu lesen: Section 12.2, Seiten 412-416 in Wooldridge (2009).

8.2. Nicht streng exogene Regressoren

Ist das Modell $y_t = \mathbf{x}_t\boldsymbol{\beta} + u_t$ dynamisch vollständig spezifiziert, gilt:

$$E[y_t|\mathbf{x}_t, y_{t-1}, \mathbf{x}_{t-1}, y_{t-2}, \ldots] = E[y_t|\mathbf{x}_t] = \mathbf{x}_t\boldsymbol{\beta}$$

und somit nach Einsetzen von $y_t = \mathbf{x}_t\boldsymbol{\beta} + u_t$:

$$E[u_t|\mathbf{x}_t] = 0$$

und

$$E[u_t|\mathbf{x}_t, u_{t-1}, \mathbf{x}_{t-1}, u_{t-2}, \ldots] = 0.$$

Durch Anwendung des Gesetzes des iterierten Erwartungswertes lassen sich beispielsweise alle verzögerten u_t's bis auf jeweils ein u_{t-j} aus der Bedingung eliminieren, so dass gilt:

$$E[u_t|u_{t-j}, \mathbf{x}_t, \mathbf{x}_{t-1}, \ldots] = 0, \quad j = 1, 2, \ldots.$$

Man beachte, dass $\mathbf{x}_t$ verzögerte y_t enthalten kann und damit verzögerte u_t's. Deshalb kann man nicht $\mathbf{x}_t$ aus der Bedingung herausintegrieren. Für $\mathbf{x}_{t-1}, \ldots$ geht das jedoch schon, da alle relevanten Regressoren bereits in $\mathbf{x}_t$ enthalten sind. Es gilt also:

$$E[u_t|u_{t-j}, \mathbf{x}_t] = 0, \quad j = 1, 2, \ldots,$$

und

$$Cov(u_t, u_{t-j}|\mathbf{x}_t) = E(u_t \cdot u_{t-j}|\mathbf{x}_t) = E\left[u_{t-j}E[u_t|u_{t-j}, \mathbf{x}_t]|\mathbf{x}_t\right] = 0$$

und die u_t's sind unkorreliert, wenn $\mathbf{x}_t$ berücksichtigt wird.

Tests auf Autokorrelation

Also lassen sich die Null- und Alternativhypothese in der Regression:

$$u_t = \gamma_0 + \gamma_1 x_{t1} + \cdots \gamma_k x_{tk} + \delta_1 u_{t-1} + \cdots + \delta_q u_{t-q} + v_t \quad (8.1)$$

formulieren durch:

$$\mathsf{H}_0 : \delta_1 = 0,\ \delta_2 = 0,\ \ldots,\ \delta_q = 0$$

versus

$$\mathsf{H}_1 : \delta_j \neq 0 \text{ für mindestens ein } j = 1, \ldots, q.$$

- Die Wahl von q hängt davon ab, welche Ordnung der Autokorrelation man unter der Alternative erwartet!

- Die gemeinsame Hypothese kann dann mit einem F-**Test** überprüft werden, wobei der F-Test nur approximativ gilt, da auch die $\hat{\boldsymbol{\beta}}$ im Regressionsmodell selbst nur asymptotisch normalverteilt sind.

- Wie im vorherigen Fall werden die unbekannten Fehler durch die Residuen ersetzt und der approximative F-Test bleibt gültig. Der kritische Wert wird mit q und $T - k - q$ Freiheitsgraden bestimmt.

- Alternativ lässt sich ein so genannter **Lagrange-Multiplikator-Test** (LM-Test) durchführen. Die Teststatistik lautet:

$$LM = (T - q)R_u^2,$$

wobei R_u^2 das Bestimmtheitsmaß der Regression (8.1) ist, jedoch auf Basis der Residuen $\hat{u}_t$, ist. Die Teststatistik wird hier nicht abgeleitet. Lagrange-Multiplikator-Tests sind ein allgemeines Testverfahren, das auf viele verschiedene Testsituationen angewendet bzw. angepasst werden kann. Der Ausgangspunkt ist immer, dass das H_1-Modell unter den Restriktionen von H_0 geschätzt wird. Da hierbei eine Optimierung unter Nebenbedingungen durchgeführt werden muss, wird der Lagrange-Multiplikator-Ansatz verwendet. Hieraus ergibt sich der Name der Teststatistik.

Im vorliegenden Fall ist der LM-Test auch als **Breusch-Godfrey Test** oder als LM-**Test auf serielle Autokorrelation** bekannt. Die Teststatistik ist asymptotisch χ^2-verteilt mit q Freiheitsgraden:

$$LM \xrightarrow{d} \chi^2(q).$$

- Sowohl für den F-Test, als auch für den LM-Test gibt es Versionen für heteroskedastische Fehler. Literaturhinweise finden sich z.B. in Wooldridge (2009).

- Werden Zeitreihen mit Saisonkomponenten verwendet, kann man gezielt einzelne Autokorrelationen prüfen, z.B. bei Quartalsdaten $u_t = \rho_4 u_{t-4} + e_t$.

Zu lesen: Section 12.2, Seiten 416-419 in Wooldridge (2009).

8.3. Verallgemeinerter KQ-Schätzer bei autokorrelierten Residuen erster Ordnung

Verallgemeinerter KQ-Schätzer (Generalized least squares, GLS)

8.3.1. GLS-Schätzer für das einfache Regressionsmodell

Ableitung des GLS-Schätzers für das einfache Regressionsmodell ohne Matrixalgebra.

Gegeben sei das einfache Regressionsmodell

$$
\begin{aligned}
y_t &= \beta_0 + \beta_1 x_t + u_t, \\
u_t &= \rho u_{t-1} + e_t, \quad e_t \sim \text{i.i.d.}(0, \sigma_e^2),
\end{aligned}
$$

wobei ρ **als bekannt vorausgesetzt wird** und x_t streng exogen ist. Die Annahme strenger Exogenität kann abgeschwächt werden, siehe Ende dieses Abschnitts.

Transformiertes Modell

Einsetzen von $u_j = y_j - \beta_0 - \beta_1 x_j$, $j = t, t-1$, in die AR(1)-Gleichung von u_t und Umformulieren ergibt:

$$\underbrace{y_t - \rho y_{t-1}}_{y_t^*} = \beta_0 - \rho\beta_0 + \beta_1(\underbrace{x_t - \rho x_{t-1}}_{x_t^*}) + \underbrace{e_t}_{u_t^*}, \quad t = 2, 3, \ldots,$$

$$y_t^* = \beta_0(1-\rho) + \beta_1 x_t^* + u_t^*, \quad t = 2, 3, \ldots. \tag{8.2}$$

Dieses transformierte Modell (quasi-Differenzen) hat unkorrelierte Fehler, da e_t i.i.d. ist, und erfüllt damit die Annahmen TS.1 bis TS.5.

Die Schätzung von (8.2) ist damit mit dem gewöhnlichen OLS-Schätzer möglich und weist die bekannten asymptotische Verteilungseigenschaften auf.

Allerdings ist der Schätzer ineffizient, da die erste Beobachtung verloren geht.

Berücksichtigung erste Beobachtung

Obwohl u_1 mit allen $e_t = u_t^*$, $t = 2, 3, \ldots, T$, unkorreliert ist, kann man nicht einfach (y_1, x_1) als (y_1^*, x_1^*) und somit u_1 als u_1^* verwenden, da

$$Var(u_1) = \sigma_e^2/(1-\rho^2) > Var(e_t).$$

Die erste Beobachtung hätte in (8.3) also eine andere Fehlervarianz als die für $t = 2, 3, \ldots$. Die dadurch entstehende Heteroskedastie kann man korrigieren, indem man $y_1 = \beta_0 + \beta_1 x_1 + u_1$ mit $\sqrt{(1-\rho^2)}$ multipliziert und erhält:

$$\underbrace{y_1\sqrt{(1-\rho^2)}}_{y_1^*} = \beta_0\sqrt{(1-\rho^2)} + \beta_1 \underbrace{x_1\sqrt{(1-\rho^2)}}_{x_1^*} + \underbrace{u_1\sqrt{(1-\rho^2)}}_{u_1^*}. \tag{8.3}$$

Damit können alle Beobachtungen verwendet werden und der GLS-Schätzer auf Basis von (8.2) und (8.3) ist BLUE.

Beachte: Das Modell enthält nun als ersten Regressor keine Konstante mehr, sondern den Vektor $\left(\sqrt{(1-\rho^2)} \;\; (1-\rho) \;\; \cdots \;\; (1-\rho)\right)'$.

Anwendbarkeit GLS-Schätzer

Die minimale Voraussetzung für die Anwendbarkeit des **GLS-Schätzers** im vorliegenden Beispiel ergibt sich aus $E[u_t^*|x_t^*] = 0$ und damit

$$Cov(x_t^*, u_t^*) = 0.$$

Einsetzen von x_t^* und $u_t^* = e_t$ ergibt für $t = 2, 3, \ldots$:

$$\begin{aligned} Cov(x_t - \rho x_{t-1}, u_t - \rho u_{t-1}) &= E[(x_t - \rho x_{t-1})(u_t - \rho u_{t-1})] \\ &= -\rho E[x_{t-1} u_t] - \rho E[x_t u_{t-1}] \end{aligned}$$

was bei strenger Exogenität erfüllt ist.
(Zeigen Sie, dass auch für $t = 1$ gilt: $Cov(x_1^*, u_1^*) = 0$.)

Abschwächung: Ist x_t stationär, gilt $E[x_t u_{t-1}] = E[x_{t+1} u_t]$, so dass

die Kovarianz Null ist, wenn

$$E[(x_{t-1} + x_{t+1})u_t] = 0.$$

Ist diese Bedingung nicht erfüllt, dann muss das Modell entweder doch noch dynamisch vollständig spezifiziert werden oder es müssen autokorrelations- (und heteroskedastie-) robuste Standardfehler verwendet werden, siehe hierzu Abschnitt 8.5.

Zu lesen: Section 12.3, Seiten 419-421 in Wooldridge (2009).

8.3.2. GLS für das allgemeine Regressionsmodell

GLS für das allgemeine Regressionsmodell in Matrixdarstellung.

Wiederholung aus Abschnitt 7.2 in **Einführung in die Ökonometrie**:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad \mathbf{u}|\mathbf{X} \sim \left(0, \sigma^2\boldsymbol{\Psi}\right), \quad \mathbf{P}'\mathbf{P} = \boldsymbol{\Psi}^{-1}, \tag{8.4}$$

$$\mathbf{P}\mathbf{y} = \mathbf{P}\mathbf{X}\boldsymbol{\beta} + \mathbf{P}\mathbf{u},$$

$$\mathbf{y}^* = \mathbf{X}^*\boldsymbol{\beta} + \mathbf{u}^*. \tag{8.5}$$

Der Vektor $\mathbf{u}^*$ ergibt sich aus den u_t^* aus dem vorigen Abschnitt.

Wie lautet $\mathbf{P}$?

Schreibt man y_1^* und $y_2^*, \ldots, y_T^*$ in Matrixform, erhält man

$$\begin{pmatrix} y_1^* \\ y_2^* \\ y_3^* \\ \vdots \\ y_T^* \end{pmatrix} = \underbrace{\begin{pmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{pmatrix}}_{\mathbf{P}} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_T \end{pmatrix} = \mathbf{P}\mathbf{y}.$$

Entsprechend ergeben sich $\mathbf{X}^* = \mathbf{P}\mathbf{X}$ und $\mathbf{u}^* = \mathbf{P}\mathbf{u}$.

Da $\mathbf{P}'\mathbf{P} = \mathbf{\Psi}^{-1}$ ergibt sich:

$$\mathbf{\Psi}^{-1} = \begin{pmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{pmatrix},$$

$$\mathbf{\Psi} = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{pmatrix}.$$

(Inversion ohne Beweis.) Interpretieren Sie jedoch $\mathbf{\Psi}$!

8.4. FGLS-Schätzer bei AR(1)-Fehlern

Anwendbarer verallgemeinerter Kleinst-Quadrate-Schäter (Feasible generalized least squares, FGLS).

Ist der Autokorrelationskoeffizient ρ des AR(1)-Prozesses in den Fehlern unbekannt, muss dieser ebenfalls geschätzt werden und der FGLS verliert seine BLUE-Eigenschaft (so wie bei einer zu schätzender Heteroskedastiefunktion (siehe **Einführung in die Ökonometrie**)).

Kann ρ konsistent geschätzt werden, lässt sich der FGLS-Schätzer verwenden. (Welche Annahmen sind zu erfüllen, damit die u_t's konsistent geschätzt werden können?)

- **FGLS-Schätzer bei autokorrelierten Residuen**

  i) Regressiere mit OLS y_t auf $\mathbf{x}_t$ und speichere die $\hat{u}_t$'s.

  ii) Schätze mit OLS ρ in $\hat{u}_t = \rho \hat{u}_{t-1} + e_t$.

  iii) Schätze Gleichung (8.4) mit (8.5), wobei ρ in $\mathbf{P}$ durch $\hat{\rho}$ ersetzt wird und $\hat{\mathbf{P}}$ statt $\mathbf{P}$ geschrieben wird. In Matrixschreibweise lautet die letzte Stufe:

$$\hat{\mathbf{P}}\mathbf{y} = \hat{\mathbf{P}}\mathbf{X}\boldsymbol{\beta} + \hat{\mathbf{P}}\mathbf{u}.$$

  Man beachte, dass der FGLS ein mehrstufiger Schätzer ist.

- **Schätzeigenschaften**: Aufgrund der Schätzung von ρ geht die BLUE-Eigenschaft verloren. Allerdings wurde gezeigt, dass asymptotisch die Schätzung von ρ keine Rolle spielt, so dass der FGLS-Schätzer.

  – asymptotisch normalverteilt ist, vorausgesetzt für Gleichung (8.5) sind alle Annahmen TS.1' bis TS.5' erfüllt. Auf (8.5) lassen sich deshalb die bisherigen t- und F-Tests anwenden. Achtung: Das obige Beispiel zeigt, dass die Annahmen TS.1' bis TS.5' für das transformierte Modell (8.5) *nicht* erfüllt sein müssen, selbst wenn TS.1' bis TS.3' für das ursprüngliche Modell erfüllt sind;

  – asymptotisch effizient ist, d.h. unter allen konsistenten Schätzern die "kleinste" Varianz aufweist.

- Im einzelnen gibt es eine Reihe von **Variationen des FGLS-Schätzers**:
  - **Cochrane-Orcutt (CO) Schätzer**: In Schritt iii) wird die erste Beobachtung weggelassen, so dass statt (8.4) die multivariate Variante von (8.2) geschätzt wird.
  - **Prais-Winsten (PW) Schätzer**: Schätzung von (8.5).
  - Iterierte Schätzer: CO bzw. PW werden mehrmals in Folge geschätzt, bis sich die Parameter kaum mehr verändern.

Die asymptotischen Schätzeigenschaften aller genannten Schätzer sind gleich, da asymptotisch die erste Beobachtung keine Rolle spielt und damit CO und PW asymptotisch effizient sind. In kleinen Stichproben unterscheiden sich die Schätzer jedoch.

- In Appendix A.15 findet sich ein R-Programm, das einen FGLS-Schätzer bei autokorrelierten Fehlern illustriert und ein entsprechender HAC-Schätzer, der in Abschnitt 8.5 eingeführt wird.

- **OLS versus FGLS**

  - Die Schätzergebnisse unterscheiden sich numerisch fast immer.

  - Die Annahmen für Konsistenz des OLS-Schätzers sind schwächer als im Fall des GLS-Schätzers (z.B. ist x_{t-1} als Regressor ausgeschlossen, siehe oben).

  - Schätzunterschiede zwischen beiden Schätzmethoden können also auf Annahmeverletzungen zurückzuführen sein. Ein Test hierfür (Hausman-Test) übersteigt das Niveau dieses Kurses.

- Sind die Annahmen für FGLS erfüllt, ist FGLS asymptotisch effizient.

- **Weitere Anmerkungen**

  - Korrektur von Autokorrelation höherer Ordnung ist mühsamer, siehe z.B. Ende von Section 12.3 in Wooldridge (2009).

  - Vermutet man sehr große Autokorrelation in den Fehlern (ρ nahe 1), schlägt Wooldridge (2009) in Section 12.4 vor, erste Differenzen zu bilden, da damit ρ gut approximiert wird.

8.5. Autokorrelations- und heteroskedastie-robuste Standardfehler für den OLS-Schätzer

- Die Annahmen für die Anwendung des FGLS-Schätzers sind nicht erfüllt.
- Die Art der Autokorrelation in den Fehlern ist unbekannt.
- Weitere Gründe siehe unten.

8.5.1. Ableitung des autokorrelations-robusten Standardfehlers für β_1 (zur Info)

Gegeben sei

$$y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t, \quad u_t \text{ stationär.}$$

Man beachte nun, dass in der OLS-Hilfsregression

$$x_{t1} = \delta_0 + \delta_2 x_{t2} + \ldots + \delta_k x_{tk} + r_t$$

die Residuen $\hat{r}_t$ mit den "Regressoren" x_{t2} bis x_{tk} unkorreliert sind. Damit lässt der Einfluss von x_{t1} auf y_t aufspalten in den x_{t1} eigenen Einfluss verkörpert durch $\hat{r}_t$ und den Einfluss der indirekt durch die x_{t2} bis x_{tk} kommt. Einsetzen ergibt:

$$y_t = \beta_0 + \beta_1\hat{\delta}_0 + \left(\beta_1\hat{\delta}_2 + \beta_2\right) x_{t2} + \ldots + \left(\beta_1\hat{\delta}_k + \beta_k\right) x_{tk} + \beta_1\hat{r}_t + u_t.$$

Da jetzt $\hat{r}_t$ mit allen anderen Regressoren in der Stichprobe unkorreliert ist, kann man gemäß den Ergebnissen des omitted variable bias (siehe Abschnitt 7.2) auch

$$y_t = \beta_1 \hat{r}_t + u_t$$

schätzen und erhält für $\hat{\beta}_1$:

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{t=1}^{T} \hat{r}_t u_t}{\sum_{t=1}^{T} \hat{r}_t^2}.$$

Damit ist die Varianz von $\hat{\beta}_1$:

$$Var(\hat{\beta}_1) = Var\left(\frac{\sum_{t=1}^{T} \hat{r}_t u_t}{\sum_{t=1}^{T} \hat{r}_t^2}\right).$$

Man kann nun mit Werkzeugen der mathematischen Statistik zeigen, dass asymptotisch, d.h. in "**großen Stichproben**" gilt:

$$Var(\hat{\beta}_1) \approx \frac{Var\left(\sum_{t=1}^{T} r_t u_t\right)}{\left(\sum_{t=1}^{T} E[r_t^2]\right)^2}. \tag{8.6}$$

Ist der Fehlerprozess $\{u_t\}$ autokorreliert, sind es auch die $a_t = r_t u_t$ und man erhält für die Varianz im Zähler (siehe auch Beginn von Abschnitt 8.1):

$$Var\left(\sum_{t=1}^{T} r_t u_t\right) = \sum_{t=1}^{T} Var(a_t) + 2 \sum_{t=1}^{T-1} \sum_{j=1}^{T-t} Cov(a_t, a_{t+j}).$$

Ersetzt man den Index durch $s = t + j$ erhält man

$$Var\left(\sum_{t=1}^{T} r_t u_t\right) = \sum_{t=1}^{T} Var(a_t) + \underbrace{2\sum_{j=1}^{T-1}\sum_{s=j+1}^{T} Cov(a_{s-j}, a_s)}$$

tritt nur bei autokorr. Fehlern auf

Da bei stationären Fehlern die Autokovarianzen für große j klein sind, approximiert man den hervorgehobenen Term, indem man nur die ersten g Autokovarianzen berücksichtigt und schätzt. Entsprechend der Ableitung heteroskedastie-robuster Standardfehler (**Einführung in die Ökonometrie**, Abschnitt 7.1) schätzt man $Cov(a_s, a_{s-j})$ durch $\hat{a}_s \hat{a}_{s-j}$. Damit erhält man folgenden Schätzer für den Zähler in (8.6):

$$\widehat{Var}\left(\sum_{t=1}^{T} r_t u_t\right) = \sum_{s=1}^{T} \hat{a}_s^2 + 2\sum_{j=1}^{g}\left(1 - \frac{j}{g+1}\right)\sum_{s=j+1}^{T} \hat{a}_{s-j}\hat{a}_s,$$

wobei der blaue Term aus schätztechnischen Gründen eingefügt wurde.

Da der Nenner in (8.6) mit dem Nenner der Standard OLS-Varianz für $\hat{\beta}_1$ übereinstimmt, lässt sich (8.6) berechnen durch:

$$\widehat{Var}(\hat{\beta}_1) = \frac{\sigma^2}{\hat{\beta}_1}\widehat{Var}\left(\sum_{t=1}^{T} r_t u_t\right).$$

8.5.2. Wahl der Anzahl g der geschätzten Autokovarianzen

- Beachte: Je größer g, desto unsicherer wird die Schätzung des Standardfehlers, wenn keine Autokorrelation in den Fehlern vorliegt, da sozusagen überflüssige Parameter geschätzt werden.
- Allgemein muss g geeignet mit der Stichprobengröße wachsen, damit in großen Stichproben auch komplizierte Autokorrelationsstrukturen erfasst werden können.

- Wird in `R` im Befehl `coeftest()` aus dem Paket `lmtext` für den Parameter `vcov` die Funktion `Newey-West()` aus dem Paket `sandwich` verwendet, wird g mit dem von Newey und West (1994) vorgeschlagenen automatischen Verfahren bestimmt.

- Siehe das `R`-Programm in Appendix A.15 für ein Beispiel.

- g muss groß sein, wenn die Fehler einem AR(1)-Prozess mit ρ nahe 1 folgen, da dann $Cov(a_s, a_{s-j})$ auch für große j groß sind!

- Wenn die Fehler einem MA(1)-Prozess folgen, funktionieren hingegen die HAC-Standardfehler hervorragend, da $g = 1$ genügen würde.

8.5.3. Heteroskedastie in den Fehlern

Die autokorrelations-robusten Varianzen sind auch heteroskedastie-robust. Sie werden deshalb als **heteroskedasticity and autocorrelation consistent** (HAC) "covariances" bezeichnet.

Warnhinweise

- In kleinen Stichproben ist der HAC-Schätzer unter Umständen ein ungenauer Schätzer der tatsächlichen Schätzvarianz!
- Auch kann die Signifikanz von Parametern von der Wahl von g abhängen.

- Benutze HAC nicht direkt, falls ρ nahe 1! Hier sollte man die Daten zuerst quasi-differenzieren und dann HAC anwenden, siehe Section 12.5, Seite 430f in Wooldridge (2009).

Zu lesen: Section 12.5 in Wooldridge (2009).

8.6. Heteroskedastie in Zeitreihen

Zum Selbststudium (nicht klausurrelevant): Section 12.6 in Wooldridge (2009).

8.7. Unendlich-verteilte Lag Modelle

Unendlich-verteilte Lag Modelle **infinite distributed lag (IDL) models** gibt es einige:

- Geometrische oder Koyck-verteilte Lags (**geometric or Koyck distributed lags**).
- Rational verteilte Lags (**rational distributed lag, RDL models**).

Diese Modelle erfordern strenge Exogenität der Regressoren und erlauben dann die Berechnung von kurz- und langfristigen Multiplikatoren wie in Abschnitt 2.1.

Zum Selbststudium (nicht klausurrelevant): Section 18.1 in Wooldridge (2009).

9. Tests zum Überprüfen der Random Walk-Hypothese: Einheitswurzeltests

9.1. Dickey-Fuller-Test

Aus Abschnitt 6.2 ist bekannt, dass Random Walks nicht stationär sind und deshalb die OLS-Schätzer der Parameter ν und ρ (bzw. θ) des AR(1)-Modells auch asymptotisch nicht normalverteilt sind:

$$y_t = \nu + \rho y_{t-1} + e_t \quad \text{bzw.} \quad \Delta y_t = \nu + \underbrace{(\rho - 1)}_{\theta} y_{t-1} + e_t.$$

Stattdessen folgen die Schätzer einer Nichtstandardverteilung und auch die gewöhnlichen t-Statistiken sind nichtstandard verteilt.

Da Dickey und Fuller 1979 diese Verteilung abgeleitet haben, wird diese häufig als **Dickey-Fuller-Verteilung** (DF-Verteilung) bezeichnet.

Dabei hängt die spezifische Form der DF-Verteilung von der Spezifikation der Alternative ab. Im folgenden darf e_t bedingt heteroskedastisch sein, jedoch muss e_t unkorreliert sein, d.h. y_t ist dynamisch vollständig spezifiziert.

- Fall A: nur Konstante, kein Zeittrend unter der Alternativhypothese.
- Fall B: Konstante und linearer Zeittrend unter der

Alternativhypothese.

Fall A: nur Konstante, kein Zeittrend unter der Alternativhypothese

$$\mathsf{H}_0 : \rho = 1 \text{ versus } \mathsf{H}_1 : \rho < 1,$$

bzw.

$$\mathsf{H}_0 : \theta = 0 \text{ versus } \mathsf{H}_1 : \theta < 0,$$

mit

$$\Delta y_t = \nu + \theta y_{t-1} + e_t$$

als Modell unter der Alternative.

Fall B: Konstante und linearer Zeittrend unter der Alternativhypothese

$$\mathsf{H}_0 : \rho = 1 \text{ versus } \mathsf{H}_1 : \rho < 1,$$

bzw.

$$\mathsf{H}_0 : \theta = 0 \text{ versus } \mathsf{H}_1 : \theta < 0,$$

mit

$$\Delta y_t = \nu + \delta t + \theta y_{t-1} + e_t$$

als Modell unter der Alternative.

Asymptotische kritische Werte für den DF-Test

Die Nullhypothese wird abgelehnt, wenn $\hat{t}_\theta < c$ ist, wobei der kritische Wert c für beide Fälle in folgenden Tabellen tabelliert ist:

Tabelle 9.1.: Asymptotische kritische Werte für Dickey-Fuller-Einheitswurzeltests, Quelle: Davidson und MacKinnon (1993), Tabelle 20.1, S. 708.

Quantil	1%	2.5%	5%	10%	97.5%
			t-Statistik		
Hypothesenpaar (A)	-3.43	-3.12	-2.86	-2.57	0.24
Hypothesenpaar (B)	-3.96	-3.66	-3.41	-3.13	-0.66

Diese Tests heißen **Dickey-Fuller Einheitswurzeltest** (kurz DF-Test).

Beachte:

- Diese **kritischen Werte sind für** $T \to \infty$ berechnet. Man kann jedoch **kritische Werte für jedes beliebige** T mit Hilfe von Monte Carlo Simulationen berechnen. Entsprechende kritische Werte werden z. B. im R-Paket `urca()` in dem Befehl `ur.df` verwendet. Dann weichen die kritischen Werte bei kleinen Stichproben von den obigen Tabellenwerten ab.

- Ist man sich unsicher (z. B. nach graphischer Analyse), ob unter der Alternative ein Trend vorliegt, so wählt man im Allgemeinen Fall B. Häufig erfolgt diese Entscheidung nach ökonomischen Gesichtspunkten. Z.B. ist es unter normalen ökonomischen Bedingungen unwahrscheinlich, dass Zinsen einen Zeittrend aufweisen.

- Die t-Statistik des Trendparameters ist nur dann asymptotisch normalverteilt, wenn $|\rho| < 1$.

- Folgt y_t einem AR-Prozess höherer Ordnung, ist im vorliegenden Fall das Modell dynamisch unvollständig spezifiziert.

  Lösungen:

  – Den Dickey-Fuller-Test mit der korrekten Anzahl an Lags ausführen (siehe nächster Punkt).

  – Z.B. den Phillips-Perron-Test verwenden, der die Autokorrelation in den Fehlern korrigiert (hier nicht behandelt).

9.2. Augmented Dickey-Fuller-Test

Liegt $\{y_t\}$ ein AR(p)-Prozess zugrunde, dann lässt sich dieser schreiben als:

- Fall A: nur Konstante, kein Zeittrend unter der Alternativhypothese:

$$\Delta y_t = \nu + \theta y_{t-1} + \alpha_1^* \Delta y_{t-1} + \ldots + \alpha_{p-1}^* \Delta y_{t-p+1} + e_t.$$

- Fall B: Konstante und linearer Zeittrend unter der Alternativhypothese:

$$\Delta y_t = \nu + \delta t + \theta y_{t-1} + \alpha_1^* \Delta y_{t-1} + \ldots + \alpha_{p-1}^* \Delta y_{t-p+1} + e_t.$$

Es gilt:

- Für den Augmented Dickey-Fuller-Test (ADF-Test) werden für $\hat{t}_\theta$

wieder die kritischen Werte aus obiger Tabelle verwendet.

- Die t-Statistiken für die α_j^* sind asymptotisch normalverteilt (sofern e_t homoskedastisch ist), d.h. approximativ t-verteilt (da die Regressoren sowohl unter H_0 als auch unter H_1 stationär sind).

- Die **korrekte Lag-Ordnung** wird am einfachsten mit dem AIC- oder dem Schwarz-Kriterium bestimmt. Dies ist sowohl unter H_0 als auch unter H_1 möglich. Man beachte, dass beim Vergleich von Modellen mit unterschiedlicher Lag-Ordnung **alle** Modelle mit den gleichen Startwerten geschätzt werden. Ansonsten ist das Selektionsergebnis in kleinen Stichproben häufig irreführend.

Im Appendix A.1 befindet sich ein R-Programm, das ADF-Tests für die Arbeitslosenraten und Inflation aus Abschnitt 1.1.1 berechnet. Im Appendix A.16 befindet sich ein R-Programm, das ADF-Tests für den

realen US-Aktienindex S&P 500 aus Abschnitt 1.1.2 berechnet.

Zu lesen: Section 18.2 in Wooldridge (2009).

10. Kointegration und Vektorfehlerkorrekturmodelle

10.1. Erinnerung: Besonderheiten bei Regressionen mit I(1)-Variablen

Vgl. Abschnitt 6.4:

- im Allgemeinen asymptotische Nichtstandardverteilung des OLS-Schätzers,

- spezifische Interpretation der Parameter: Kointegrationsbeziehungen zwischen stochastischen und ggf. deterministischen Trends,
- Abhängigkeit der Schätzeigenschaften von der Formulierung der Regression in Niveaus oder ersten Differenzen,
- Gefahr einer Scheinregressionen (spurious regression).

10.2. Kointegrationsbeziehungen und Fehlerkorrekturmodelle für zwei I(1)-Variablen

Der einfachste Fall einer Kointegrationsbeziehung liegt vor für zwei I(1)-Variablen y_t und x_t, so dass

$$\Delta y_t = \mu_1 + \tilde{y}_t,$$
$$\Delta x_t = \delta_1 + \tilde{x}_t.$$

Vgl. Fall I in Abschnitt 6.4.1, Gleichung (6.5):

$$y_t = \beta_0 + \beta_1 x_t + u_t, \quad u_t \sim WN(0, \sigma^2),$$

wobei hier $\mu_1 = \beta_1 \delta_1$ (vgl. Fall IV in Abschnitt 6.4.4).

Die Annahme, dass u_t weißes Rauschen ist, kann abgeschwächt werden zu $u_t \sim$ I(0). Dann ist es möglich, dass u_t weitere I(0)-Regressoren enthält. Alternativ kann man diese explizit modellieren.

Einfluss von kurzfristiger Dynamik und anderen I(0)-Regressoren

Die Kurzfristdynamik wird in einem Fehlerkorrektur-Modell wie Gleichung (6.6) abgebildet.

Eine erweiterte Variante einer derartigen Spezifikation mit zusätzlichen Regressoren ist z.B.:

$$\begin{aligned}\Delta y_t = \nu + \alpha\left(y_{t-1} - \beta_1 x_{t-1}\right) \\ + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + \alpha_1 \Delta y_{t-1} + v_t.\end{aligned}$$

Darstellung in Trendvariablen

Man kann die Gleichung auch vollständig in den Trendvariablen schreiben, indem man $\Delta y_t = \mu_1 + \tilde{y}_t$, $\Delta x_t = \delta_1 + \tilde{x}_t$ einsetzt. Dies ergibt:

$$\begin{aligned} y_t &= \nu + \alpha y_{t-1} - \alpha\beta_1 x_{t-1} + \gamma_0 x_t - \gamma_0 x_{t-1} + \gamma_1 x_{t-1} - \gamma_1 x_{t-2} \\ &\quad + \alpha_1 y_{t-1} - \alpha_1 y_{t-2} + v_t \\ &= \tilde{\beta}_0 + \tilde{\beta}_1 x_t + \tilde{\beta}_2 x_{t-1} + \tilde{\beta}_3 x_{t-2} + \tilde{\beta}_4 y_{t-1} + \tilde{\beta}_5 y_{t-2} + v_t. \end{aligned}$$

- Man erhält also ein AR(2)-Modell ergänzt um einen finite distributed lag (FDL)-Teil.
- Die Kointegrationsbeziehung ist jetzt nicht mehr direkt sichtbar, wirkt sich jedoch dadurch aus, dass die Gleichung weniger unbekannte Parameter als erklärende Variablen enthält (5 bzw. 6 inkl. Scheinregressor).

- Dies ist die allgemeine Form eines **Fehlerkorrekturmodells**.

- Modelliert man zusätzlich ein entsprechendes Fehlerkorrekturmodell für Δx_t gemeinsam mit dem Fehlerkorrekturmodell für Δy_t, erhält man ein **Vektorfehlerkorrekturmodell**. Siehe Veranstaltung **Quantitative Wirtschaftsforschung I** für Details und Anwendungen.

10.3. Kointegrationsbeziehungen mit mehreren I(1)-Variablen

Die Kointegrationsbeziehung lautet dann:

$$y_t = \mathbf{x}_t\boldsymbol{\beta} + u_t, \tag{10.1}$$

wobei $\mathbf{x}_t$ einen Scheinregressor und ggf. einen linearen Trend (Fall IV in Abschnitt 6.4.4) enthalten kann.

Beachte:

- Bei Vorliegen von mehr als zwei I(1)-Variablen kann es mehr als eine Kointegrationsbeziehung geben.
- Um ggf. mehr als eine Kointegrationsbeziehung explizit zu modellieren und zu schätzen müssen jedoch Vektorfehlerkorrekturmodelle verwendet werden.

10.4. Schätzung der Parameter in Kointegrationsbeziehungen

Der OLS-Schätzer für $\boldsymbol{\beta}$ in Gleichung (10.1) ist asymptotisch nicht normalverteilt, jedoch konsistent.

10.4.1. Strenge Exogenität der Regressoren $\mathbf{x}_t$

Sind die $\mathbf{x}_t$ streng exogen (und Gleichung (10.1) korrekt spezifiziert), dann gelten Annahmen TS.1 bis TS.3.

- Gelten darüber hinaus TS.4 bis TS.6, d.h. sind die **Fehler u_t gegeben $\mathbf{X}$ normalverteilt**, dann ist der OLS-Schätzer **gegeben die $\mathbf{X}$** exakt normalverteilt!

- Sind die Fehler normalverteilt, aber autokorreliert und heteroskedastisch, können die HAC-Standardfehler verwendet werden oder der FGLS-Schätzer. Dann gilt jedoch nur noch eine asymptotische Normalverteilung und t-Statistiken haben eine asymptotische Standardnormalverteilung.

- Ist die Fehlerverteilung unbekannt, dann kann gezeigt werden, dass der OLS-Schätzer von $\boldsymbol{\beta}$ asymptotisch normalverteilt ist und die t-Statistiken asymptotisch standardnormalverteilt sind.

- In allen genannten Fällen funktionieren damit F-Tests exakt oder zumindest asymptotisch.

10.4.2. Regressoren x_t nicht streng exogen

Herleitung nur für den skalaren Fall $\mathbf{x}_t = x_t$.

Was tun, wenn x_t nicht streng exogen ist? Man kann x_t approximativ streng exogen machen. Strenge Exogenität erfordert, dass nicht nur $Cov(u_t, x_{t-j}) = 0$, $j = 1, 2, \ldots$, sondern auch $Cov(u_t, x_{t+j}) = 0$, $j = 1, 2, \ldots$ gilt.

Man kann dies erreichen, indem man einen "neuen" Fehler e_t erzeugt, für den diese Anforderung erfüllt ist, indem man eine Regression von u_t auf alle vergangenen und zukünftigen Δx_t durchführt:

$$u_t = \eta + \phi_0 \Delta x_t + \phi_1 \Delta x_{t-1} + \phi_2 \Delta x_{t-2} + \ldots + \phi_{t-1} \Delta x_{t-(t-1)} + \phi_{-1} \Delta x_{t+1} + \ldots + \phi_{-(T-t)} \Delta x_{t+(T-t)} + e_t.$$

Man beachte, dass man erste Differenzen nimmt, da diese I(0) sind.

Dynamic OLS (DOLS) estimator

In der Praxis kann man obige Regression nicht schätzen. Da aber bei stationären Zeitreihen Kovarianzen für große Abstände j nahe Null sind, kann man obige Regression gut approximieren, indem man nur p Leads und Lags verwendet:

$$\begin{aligned} u_t = \eta + \phi_0 \Delta x_t + \phi_1 \Delta x_{t-1} + \ldots + \phi_p \Delta x_{t-p} \\ + \phi_{-1} \Delta x_{t+1} + \ldots + \phi_{-p} \Delta x_{t+p} + e_t \end{aligned}$$

und in Gleichung (10.1) einsetzt. Man erhält den sogenannten **leads and lags estimator** oder **dynamic OLS (DOLS) estimator**:

$$\begin{aligned} y_t = \beta_0 + \beta_1 x_t \\ + \eta + \phi_0 \Delta x_t + \phi_1 \Delta x_{t-1} + \ldots + \phi_p \Delta x_{t-p} \\ + \phi_{-1} \Delta x_{t+1} + \ldots + \phi_{-p} \Delta x_{t+p} + e_t. \end{aligned} \tag{10.2}$$

Beachte:

- Es müssen zusätzliche Parameter geschätzt werden. Damit ist der leads and lags Schätzer in Gleichung (10.2) nicht effizient, aber eine einfache Methode.

- Sind die Fehler e_t autokorreliert und heteroskedastisch, müssen wieder HAC-Standardfehler berechnet werden oder der FGLS-Schätzer verwendet werden.

- Man kann ihn auch im Fall mehrerer I(1)-Regressoren anwenden. Allerdings wird dann die Notation etwas komplizierter.

- Im Appendix A.1 befindet sich ein R-Programm, das eine DOLS-Schätzung einer möglichen Kointegrationsbeziehung zwischen Arbeitslosenraten und Inflation aus Abschnitt 1.1.1 berechnet.

10.5. Kointegrationstests

Die Schätzung des Kointegrationsvektors setzt voraus, dass eine Kointegrationsbeziehung existiert. Sonst tritt das Problem der **Scheinregression** (Fall III in Abschnitt 6.4.3) auf.

Einfacher Fall: x_t skalar

Gegeben seien zwei I(1)-Variablen y_t und x_t. Dann liegt zwischen y_t und x_t eine Kointegrationsbeziehung vor, wenn in

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

der Fehlerprozess $u_t \sim$ I(0) ist.

Koeffizienten β_0 und β_1 bekannt

Kann u_t direkt beobachtet werden, z.B. wenn β_0 und β_1 bekannt sind, kann man mit dem ADF-**Test** den Fehlerprozess $\{u_t\}$ auf das Vorliegen einer Einheitswurzel testen.

Die ADF-Testgleichung lautet im Fall A (nur Konstante, kein Trend)

$$\Delta u_t = \nu + \theta u_{t-1} + \gamma_1 \Delta u_{t-1} + \ldots + \gamma_{p-1} \Delta u_{t-p+1} + e_t$$

Das Hypothesenpaar lautet:

$\mathrm{H}_0 : \theta = 0 \Leftrightarrow$ "keine Kointegrationsbeziehung bzw. Scheinregression"

versus

$\mathrm{H}_1 : \theta < 0$ bzw. "Vorliegen einer Kointegrationsbeziehung".

Die kritischen Werte finden sich in der Tabelle für ADF-Tests in Kapitel 9.

Man beachte, dass die Zahl der Lags für Δu_t korrekt spezifiziert werden!

Koeffizienten β_0 und β_1 unbekannt

Sind die Parameter β_0 und β_1 unbekannt und müssen geschätzt werden, müssen statt der Fehler die Residuen $\hat{u}_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t$ verwendet werden.

Unter H_0 "Random Walk in u_t" liegt eine Scheinregression vor und die OLS-Schätzer für β_0, β_1 sind nicht konsistent (da die Varianz nicht gegen Null geht). Es kann gezeigt werden, dass für die t-Statistik asym-

ptotisch eine Variante der Dickey-Fuller Verteilung gilt. Die kritischen Werte für Fall A und Fall B sind jedoch im Vergleich kleiner, da unter H_0 die OLS-Residuen häufig einem I(0)-Prozess ähnlich sind, vgl. Tabelle 10.1.

Man beachte, dass auch hier die Anzahl der Lags in der ADF-Testgleichung für $\Delta\hat{u}_t$ korrekt bestimmt wird.

Im Appendix A.1 befindet sich ein R-Programm, das einen einfachen Kointegrationstest auf Basis des ADF-Tests für mögliche Kointegration zwischen Arbeitslosenraten und Inflation aus Abschnitt 1.1.1 berechnet.

Tabelle 10.1.: Asymptotische kritische Werte für Kointegrationstests auf Basis einer Einzelgleichung, Quelle: Davidson und MacKinnon (1993), Table 20.2, S. 722.

Quantil	1%	2.5%	5%	10%	97.5
Zahl der $I(1)$-Regressoren: 1					
Hypothesenpaar (A)	-3.90	-3.59	-3.34	-3.04	-0.30
Hypothesenpaar (B)	-4.32	-4.03	-3.78	-3.50	-1.03
Zahl der $I(1)$-Regressoren: 2					
Hypothesenpaar (A)	-4.29	-4.00	-3.74	-3.45	-0.85
Hypothesenpaar (B)	-4.66	-4.37	-4.12	-3.84	-1.39
Zahl der $I(1)$-Regressoren: 3					
Hypothesenpaar (A)	-4.64	-4.35	-4.10	-3.81	-1.30
Hypothesenpaar (B)	-4.97	-4.68	-4.43	-4.15	-1.73
Zahl der $I(1)$-Regressoren: 4					
Hypothesenpaar (A)	-4.96	-4.66	-4.42	-4.13	-1.68
Hypothesenpaar (B)	-5.25	-4.96	-4.72	-4.43	-2.04
Zahl der $I(1)$-Regressoren: 5					
Hypothesenpaar (A)	-5.25	-4.96	-4.71	-4.42	-2.01
Hypothesenpaar (B)	-5.52	-5.23	-4.98	-4.70	-2.32

Vektorieller Regressor $\mathbf{x}_t$

Dieses Vorgehen funktioniert auch, falls x_t durch einen Vektor $\mathbf{x}_t$ ersetzt wird. Allerdings ändern sich dann die kritischen Werte in der Tabelle.

Zu lesen: Sections 18.3, 18.4 in Wooldridge (2009).

11. Vorhersagen (Teil II)

11.1. Allgemeine Überlegungen

Erinnerung: Abschnitt 6.1 Vorhersagen (Teil I).

11.1.1. Einige Definitionen

- Man unterscheidet **Einschritt-Vorhersagen** (one-step-ahead forecasts) und **Mehrschritt-Vorhersagen** (multi-step-ahead forecasts). Im ersten Fall wird die Zufallsvariable y_{t+1} für die nächste Periode gegeben Information bis zum Zeitpunkt t vorhergesagt, im zweiten Fall die Zufallsvariable y_{t+h}, die h Perioden in der Zukunft liegt.

- Eine Zufallsvariable y_{t+h} ist bzgl. der Informationsmenge I_t und des bedingten Erwartungswertes **nicht h Perioden in die Zukunft vorhersagbar**, wenn $E[y_{t+h}|I_t] = E[y_{t+h}]$ gilt, d.h. der bedingte Erwartungswert gleich dem unbedingten Erwartungswert ist. Die Kenntnis der Bedingung führt also zu keiner Verbesserung der Vorhersage.

- Ein stochastischer Prozess heißt **Martingal**, wenn gilt

$$E[y_{t+1}|y_t, y_{t-1}, \ldots] = y_t, \quad \text{für alle } t \geq 0.$$

Im Gegensatz zum Random Walk kann die Fehlervarianz homo- oder heteroskedastisch sein. Deshalb ist jeder Random Walk ein Martingal, aber nicht umgekehrt.

11.1.2. Prognose mit Zeitreihen-Modellen

- **Statische Modelle**, z.B.

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

sind für Vorhersagezwecke i.a. nicht gut geeignet, da zur Prognose von y_{t+h} der Regressor x_{t+h} bekannt sein muss, was selten der Fall ist. Bei unbekanntem x_{t+h} muss der Regressor selbst vorhergesagt

werden, also

$$E[y_{t+h}|I_t] = \beta_0 + \beta_1 E[x_{t+h}|I_t]$$

berechnet werden.

- **Alternative Prognose-Modelle**
  - AR(p)-Modelle,
  - Dynamische Regressionsmodelle,
  - Fehlerkorrekturmodelle,
  - VAR-Modelle.
  - Hinweis: Soweit mit Regressoren mit ähnlichen Problemen wie das statische Regressionsmodell. Vorteile reiner Zeitreihenmodelle für Prognosezwecke.

11.2. Vorhersage mit univariaten autoregressiven Modellen

Probleme mit der Vorhersage exogener Variablen werden vermieden, wenn man **autoregressive Modelle** verwendet.

11.2.1. AR(1)-Modelle

Vgl. Abschnitt 6.1.

AR(1)-Modell, gegenüber der üblichen Schreibweise eine Periode vordatiert:

$$y_{t+1} = \nu + \alpha y_t + u_{t+1}, \quad u_t \sim \text{i.i.d.}(0, \sigma^2).$$

11.2.2. Vorhersage bei bekannten Modellparametern

- Sind die Modellparameter bekannt, ergeben sich

$$
\begin{aligned}
E[y_{t+1}|y_t, y_{t-1}, \ldots] &= \nu + \alpha y_t, \\
E[y_{t+2}|y_t, y_{t-1}, \ldots] &= \nu + \alpha E[y_{t+1}|y_t, y_{t-1}, \ldots] \\
&= \nu + \alpha(\nu + \alpha y_t) \\
&= \nu(1+\alpha) + \alpha^2 y_t, \\
&\vdots \\
E[y_{t+h}|y_t, y_{t-1}, \ldots] &= \nu(1 + \alpha + \ldots + \alpha^{h-1}) + \alpha^h y_t \\
&\quad (= E[y_{t+h}|y_t]).
\end{aligned}
$$

- Der h-**Schritt Vorhersagefehler** bei bekannten Modellparametern ergibt sich, indem man die "Lösung" eines AR(1)-Prozesses (siehe Abschnitt 4.2 für den Fall ohne Konstante) einsetzt:

$$
\begin{aligned}
e_{t+h} &= y_{t+h} - E[y_{t+h}|y_t, \ldots] \\
&= \nu(1 + \alpha + \ldots + \alpha^{h-1}) + \alpha^h y_t + \sum_{j=0}^{h-1} \alpha^j u_{t+h-j} \\
&\quad - \left[\nu(1 + \alpha + \ldots + \alpha^{h-1}) + \alpha^h y_t\right] \\
&= \sum_{j=0}^{h-1} \alpha^j u_{t+h-j}.
\end{aligned}
$$

11.2.3. Vorhersagefehlervarianz

- Die **Vorhersagefehlervarianz** lässt sich daraus leicht errechnen (vgl. Abschnitt 4.2.3):

$$Var\left(e_{t+h}\right) = \sigma^2 \left(1 + \alpha^2 + \alpha^4 + \ldots + \alpha^{2h-2}\right).$$

- Die Vorhersagefehlervarianz konvergiert mit zunehmendem Vorhersagehorizont h gegen die Varianz des AR(1)-Prozesses, wenn $|\alpha| < 1$.

- Ist der AR(1)-Prozess ein **Random Walk mit oder ohne Drift** ist die **Vorhersagefehlervarianz** (bei bekanntem Trend) $Var(e_{t+h}) = \sigma^2 h$ und **wächst mit dem Vorhersagehorizont** an.

- *Für die Schätzung der Vorhersagevarianz ist es damit absolut entscheidend, ob ein Random Walk vorliegt oder nicht! Hier liegt eine weitere zentrale Bedeutung für die Anwendung der ADF-Tests!*

11.2.4. Intervallvorhersage

Die Berechnung des bedingten Erwartungswerts ist ein Beispiel einer **Punktvorhersage**.

Die Berechnung eines Vorhersageintervalls ist ein Beispiel einer **Intervallvorhersage**.

- Sind die **Fehler normalverteilt**, ist e_{t+h} ebenfalls normalverteilt und man kann ein Vorhersageintervall mit Konfidenzniveau $(1-a)\%$

berechnen durch:

$$[E[y_{t+h}|y_t] \pm z_{a/2}\sqrt{Var\left(e_{t+h}\right)}].$$

- Sind die Fehler **nicht exakt normalverteilt**, dann kann das obige Vorhersageintervall approximativ verwendet werden, wobei die Approximationsqualität umso schlechter ist, je "weiter weg" die Fehlerverteilung von der Normalverteilung ist.

11.2.5. Unbekannte Modellparameter

Bei **unbekannten** Modellparametern ersetzt man diese durch OLS-Schätzer.

- Die Vorhersage lautet dann:

$$\widehat{E}[y_{t+h}|y_t, \ldots] = \hat{\nu}(1 + \hat{\alpha} + \ldots + \hat{\alpha}^{h-1}) + \hat{\alpha}^h y_t.$$

- Da zusätzlich die Schätzunsicherheit hinzukommt, wird
  - die Vorhersagevarianz größer
  - und das Vorhersageintervall breiter (ohne Formeln).

11.2.6. AR(p)-Modelle

- Für AR-Modelle mit höherer Ordnung p funktioniert im Prinzip alles genauso. "Lediglich" die Formeln werden komplizierter.
- Beispiel: Zwei-Perioden-Prognose mit einem AR(2)-Modell (bei be-

kannten AR-Parametern):

$$\begin{aligned} y_{t+1} &= \nu + \alpha_1 y_t + \alpha_2 y_{t-1} + u_{t+1}, \\ E[y_{t+1}|y_t, y_{t-1}] &= \nu + \alpha_1 y_t + \alpha_2 y_{t-1}, \\ E[y_{t+2}|y_t, y_{t-1}] &= \nu + \alpha_1 E[y_{t+1}|y_t, y_{t-1}] + \alpha_2 y_t. \end{aligned}$$

11.2.7. Weitere Anmerkungen

- **Modelle mit linearem Trend**
  Für Vorhersagen mit großem Vorhersagehorizont ist zu beachten, dass der lineare Trend dann nicht mehr gültig sein könnte, z.B. aufgrund eines Strukturbruchs. Damit würden die Punkt-, also auch Intervallvorhersagen, u.U. stark verzerrt! Diese Unsicherheiten werden bei einem Random Walk mit Drift implizit mitberücksichtigt!

- **Modelle mit logarithmierten Variablen**
  Zur Berechnung der Vorhersagen der Niveau- (level-) Variablen, siehe z.B. Section 18.5, Seiten 655-656 in Wooldridge (2009).

11.3. Dynamische Regressionsmodelle

Hier können sowohl verzögerte endogene Variablen als auch verzögerte exogene Variablen als Regressoren verwendet werden, z.B.

$$y_t = \nu + \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} + \beta_1 x_{t-1} + \ldots + \beta_m x_{t-m} + u_t.$$

Man beachte, dass für h-Schritt Vorhersagen mit $h \geq 2$ wiederum einige x vorhergesagt werden müssen.

11.4. Vektorautoregressive Modelle

Kann man auch für x_t ein dynamisches Regressionsmodell mit verzögerten x und y spezifizieren, dann lassen sich beide Gleichungen gemeinsam modellieren und man erhält ein **vektorautoregressives**

(VAR) Modell.

Beispiel: VAR(1)-Modell

$$y_t = \nu_1 + \alpha_{11} y_{t-1} + \alpha_{12} x_{t-1} + u_t,$$
$$x_t = \nu_2 + \alpha_{21} y_{t-1} + \alpha_{22} x_{t-1} + v_t,$$

bzw. in Vektorschreibweise

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix}.$$

Mehr dazu im Kurs **Quantitative Wirtschaftsforschung I**.

11.5. Vektorfehlerkorrekturmodelle

Häufig lassen sich mehrere Fehlerkorrekturmodelle in Vektordarstellung bringen und man erhält **Vektorfehlerkorrekturmodelle (VECM) Modelle**.

VAR- und VECM-Modelle sind in der modernen Makroökonometrie ein Standardwerkzeug!

Mehr dazu im Kurs **Quantitative Wirtschaftsforschung I**.

Zu lesen: Section 18.5 in Wooldridge (2009).

A. R-Programme für die empirischen Beispiele

Hinweise:
Die Daten sind als Zip-Datei über die Homepage des Kurses verfügbar. Das Passwort wird in der Vorlesung bekanntgegeben. Die Zip-Datei muss in dem Verzeichnis entpackt werden, in dem die folgenden R-Programme gespeichert werden. Damit die Daten-Dateien von den R-Programmen gefunden werden, muss das Arbeitsverzeichnis auf das Verzeichnis gesetzt werden, in dem die folgenden R-Programme abgespeichert wurden.

A.1. Phillips-Kurve Deutschland

Verwendung:

- Deskriptiv in 1.1.1,
- Statisch in 1.1.1,
- Statisch mit Shift-Dummys in 2.5.1,
- ADF-Tests in 9.2,
- DOLS-Schätzung in 10.4.2,
- Kointegrationstest in 10.5.

Daten: `BBDL1.A.DE1.N.UNE.UBA000.A0000.A01.D00.0.R00.L.xlsx`,

BBDP1.A.DE.N.VPI.C.A00000.I20.L.xlsx, ursprüngliche Quelle: siehe Abbildung 1.1.

R-Programm:

```
################################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 1
################################################################################
#   Folien 8ff, 67, Jährliche Arbeitslosen- und Inflationsraten für Deutschland
#   Programm erstellt Zeitreihe der Daten, Graphik der Phillipskurve und
#   berechnet eine einfache lineare Regression.
#
#   Für historische Daten:
#   Die Datendateien "Phillips-Kurve_2013.xlsx" und "Phillips-Kurve_2020.xlsx"
#   müssen in dem Unterverzeichnis "Daten" liegen.
#   Für aktuelle Daten:
#   Die Dateien BBDL1.A.DE1.N.UNE.UBA000.A0000.A01.D00.0.R00.L.xlsx und
#               BBDP1.A.DE.N.VPI.C.A00000.I20.L.xlsx
#   müssen im Unterverzeichnis "Daten" liegen.
#
#   Ersteller: RT, 2019_04_25, 2021_04_15, 2025_04_26 (Aktualisierung Daten),
#               2025_07_21 Unit-Root-Tests, Kointegrationstest, DOLS-Schätzung
#               2026_02_22 Aktualisierung Daten
# ------------------------------------------------------------------------------

# ------------------------------------------------------------------------------
# Vorspann und Definitionen
# ------------------------------------------------------------------------------
```

```
# Lege Verzeichnis fest, in dem R-Programme stehen
#setwd("")
#oder
# verwende in RStudio -> Session -> Set Working Directory -> To Source File Loc.

# die beiden folgenden Variablen geben den Beginn der Daten für den Scatterplot
# und das Jahr an, nachdem die Farbe gewechselt wird.
year_begin <- 1952
year_switch <- 1970

# Logische Variable, die festlegt, ob von Graphiken PDF-Dateien erzeugt werden.
print_pdf <- TRUE       # Werte: TRUE, FALSE


# Installiere benötigte R packages und lade diese
    # Package zum Lesen von Excel-Dateien, das kein Java benötigt
if (!require(readxl)) install.packages("readxl")
library(readxl)

    # Package für spezielle Dataframes für Zeitreihendaten
    # Im Gegensatz zu ts kann zoo auch tägliche Daten handhaben
if (!require(zoo)) install.packages("zoo")
library(zoo)

# Lade Funktion, die Modellselektionskriterien wie in EViews berechnet
source("SelectCritEViews.R")

# ------------------------------------------------------------------------
# Lese Daten aus Excel-Datei ein, erstelle zoo Dataframe und berechne Inflation
```

```
# ------------------------------------------------------------------------

    # read_excel erfordert package readxl
    # data.frame erstellt Dataframe ohne Zeitangaben
        # für Daten bis 2013
# data_phillips_de <- data.frame(read_excel(path = "Daten/Phillips-Kurve_2013.xlsx",
#                                           range = "Daten für Phillipskurve!A8:E72",
#                                           col_names = TRUE))

#         # für Daten bis 2021
# data_phillips_de <- data.frame(read_excel(path = "Daten/Phillips-Kurve_2021.xlsx",
#                                           range = "Daten für Phillipskurve!A7:C77",
#                                           col_names = c("Jahr", "AL_DE",
#                                                 "P_DE")))

# für Daten bis 2024
  # Arbeitslosigkeit, Daten von Deutscher Bundesbank, Details siehe Kopf Exceldatei,
  # zum Aktualisieren gehe auf www.bundesbank.de
  # -> Statistiken -> Zeitreihen-Datenbanken -> Konjunktur und Preise
  # -> Beschäftigung und Arbeitsmarkt -> Lange Reihen
data_AL_de <- data.frame(read_excel(path = "Daten/BBDL1.A.DE1.N.UNE.UBA000.A0000.A01.D00.0.R00.L.xlsx",
                                    range = "BBDL1.A.DE1.N.UNE.UBA000.A0000.!A11:B86",
                                    col_names = c("Jahr", "AL_DE")))
  # Preisindex, Daten von Deutscher Bundesbank, Details siehe Kopf Exceldatei
  # zum Aktualisieren gehe auf www.bundesbank.de
  # -> Statistiken -> Zeitreihen-Datenbanken -> Konjunktur und Preise
  # -> Nationale Erzeuger- und Verbraucherpreise -> Nationaler Verbraucherpreisindex -> Lange Reihen
data_P_de <- data.frame(read_excel(path = "Daten/BBDP1.A.DE.N.VPI.C.A00000.I20.L.xlsx",
                                   range = "BBDP1.A.DE.N.VPI.C.A00000.I20.L!A12:B87",
                                   col_names = c("Jahr", "P_DE")))
```

```
data_phillips_de <- merge(data_AL_de, data_P_de)

    # Lese Jahr der ersten Beobachtungen aus
year_beg         <- data_phillips_de[1,1]
    # Lese Jahr der letzten Beobachtung aus
year_end         <- data_phillips_de[dim(data_phillips_de)[1],1]
    # Erstelle Dataframe mit Zeitangaben. ts ist standardmäßig in R verfügbar
data_phillips_de_ts <- ts(data = data_phillips_de[,2:dim(data_phillips_de)[2]],
                          start = year_beg, end = year_end)
    # Erstelle zoo-Dataframe. Dieser ist flexibler in der Bearbeitung von Stichproben
data_phillips_de_zoo <- zoo(data_phillips_de_ts, frequency = 1 )

# Berechne Inflationsraten
    # Beachte: im diff()-Befehl geben positive Werte bei lag Differenzen wie gewohnt an
    # Beachte: im lag()-Befehl werden gewohnte lags durch negative Werte bei k angegeben
    # Beachte die Zeitreihen aus diff() und lag() sind unterschiedlich lang. Aufgrund
    # der zeitlichen Indexierung bei zoo-Objekten werden die passenden Zeitperioden
    # zusammengefügt und die übrigen weggelassen.
INFL_DE <- diff(data_phillips_de_zoo$P_DE, lag=1) /
               lag(data_phillips_de_zoo[,"P_DE"], k=-1) * 100
    # Füge die neue "zoo"-Zeitreihe dem bisherigen zoo-Dataframe hinzu
data_phillips_de_zoo <- merge(data_phillips_de_zoo, INFL_DE)

# ------------------------------------------------------------------------------
# Erstelle Graphik der Zeitreihen für Inflation und Arbeitslosigkeit
# ------------------------------------------------------------------------------
if (print_pdf) pdf("gr_phillips.pdf", 6, 4)
plot(window(data_phillips_de_zoo[,c("AL_DE","INFL_DE")], start = year_beg,
            end = year_end),
```

```
     plot.type = "single", col = c("blue", "red"), ylab = "", xlab = "Jahr")
grid()
legend("topleft", c("AL_DE", "INFL_DE"), lty = 1, col = c("blue", "red"),
       bty = "n")
if (print_pdf) dev.off()

# ------------------------------------------------------------------------
# Erstelle Scatterplot für Inflation und Arbeitslosigkeit
# ------------------------------------------------------------------------
if (print_pdf) pdf("gr_phillips_scatter.pdf", 6, 4)
INFL_DE_range <- range(window(data_phillips_de_zoo[,"INFL_DE"], start = year_begin,
                        end = year_end), na.rm = TRUE)
plot(window(data_phillips_de_zoo[,"AL_DE"], start = year_begin, end = year_switch),
     window(data_phillips_de_zoo[,"INFL_DE"], start = year_begin, end = year_switch),
     ylim = INFL_DE_range, xlab = "Arbeitslosigkeit", ylab = "Inflation",
     type="l", col = "blue", main = "Phillips-Kurve")
#par(new = T)
lines(window(data_phillips_de_zoo[,"AL_DE"], start = year_switch + 1, end = year_end),
     window(data_phillips_de_zoo[,"INFL_DE"], start = year_switch + 1, end = year_end),
     ylim = INFL_DE_range, type="l", col = "red")
legend("topright",
       c(paste(year_begin, " - ", year_switch), paste(year_switch+1, " - ", year_end)),
       bty = "n", col = c("blue", "red"), lty = 1)
if (print_pdf) dev.off()

# ------------------------------------------------------------------------
# Berechne KQ-Schätzer für Regressionsgleichung einer einfachen Phillipskurve
# ------------------------------------------------------------------------
# für blauen Teil
phillips_de_lm <- lm(INFL_DE ~ AL_DE,
```

```
                    data=window(data_phillips_de_zoo, start = year_begin,
                    end = year_switch))
summary(phillips_de_lm)
SelectCritEViews(phillips_de_lm)

# für roten Teil
phillips_de_lm_red <- lm(INFL_DE ~ AL_DE,
                         data=window(data_phillips_de_zoo, start = year_switch+1,
                                     end = year_end))
summary(phillips_de_lm_red)
SelectCritEViews(phillips_de_lm_red)


# für blauen und roten Teil
phillips_de_lm_all <- lm(INFL_DE ~ AL_DE,
                    data=window(data_phillips_de_zoo, start = year_begin,
                    end = year_end))
summary(phillips_de_lm_all)
SelectCritEViews(phillips_de_lm_all)

# ------------------------------------------------------------------------------
# KQ-Schätzer für einfache Phillipskurve mit Strukturbruch, S. 67
# ------------------------------------------------------------------------------

    # construct dummy variables for two different periods
sev_eight    <- as.numeric(index(data_phillips_de_zoo) > 1970) *
                as.numeric(index(data_phillips_de_zoo) < 1991)
nin_zero     <- as.numeric(index(data_phillips_de_zoo) > 1990)
    # add constructed dummy variables to zoo dataframe
data_phillips_de_zoo <- merge(data_phillips_de_zoo, sev_eight,nin_zero)
```

```
    # run least squares regression
phillips_de_str_lm <- lm(INFL_DE ~ AL_DE + sev_eight + nin_zero,
                         data= window(data_phillips_de_zoo, start=1950, end=2024))
summary(phillips_de_str_lm)
SelectCritEViews(phillips_de_str_lm)

# ------------------------------------------------------------------------
#   Einheitswurzeltests / Unit Root Tests - Kointegrationstest, -schätzung
# ------------------------------------------------------------------------
# ergänzt am 21. Juli 2025

if (!require(urca)) install.packages("urca")
library(urca)

# ------------------------------------------------------------------------
#   Definiere Parameter für ADF-Tests
# ------------------------------------------------------------------------
# im Befehl in ur.df (R-Paket urca) verwendet:
adf_determ    <- "trend"    # Auswahl:
# "none": keine Konstante, kein Trend,
# "drift": nur Konstante, kein Trend
# (Fall A in Folien, S. 280),
# "trend": Konstante und Trend
# (Fall B in Folien, S. 281)
adf_maxlags   <- 5          # maximale Anzahl an Lags in Niveaus
# bei automatischer Wahl der
# Lagordnung beim ADF-Testen
adf_selcrit   <- "BIC"      # Lagauswahlkriterium: "AIC", "BIC"
# falls "Fixed" bestimmt adf_maxlags die AR-Ordnung
# der Testgleichung
```

```
p              <- 5          # AR-Ordnung
q              <- 0          # MA-Ordnung - currently only 0 allowed

# ------------------------------------------------------------------------
#           Einheitswurzeltests / Unit Root Tests
# ------------------------------------------------------------------------
summary(ur.df(INFL_DE, type = adf_determ,
              lags = adf_maxlags, selectlags = adf_selcrit))

summary(ur.df(data_phillips_de_zoo$AL_DE, type = adf_determ,
              lags = adf_maxlags, selectlags = adf_selcrit))

# ------------------------------------------------------------------------
#           Kointegrationstest via Unit Root Test
# ------------------------------------------------------------------------
# für blauen und roten Teil
phillips_de_lm_all <- lm(INFL_DE ~ AL_DE,
                         data=window(data_phillips_de_zoo, start = year_begin,
                                     end = year_end))
summary(phillips_de_lm_all)
SelectCritEViews(phillips_de_lm_all)

phillips_de_lm_all_resid <- residuals(phillips_de_lm_all)
plot(phillips_de_lm_all_resid, type = "l")

phillips_de_lm_all_resid.adf    <- ur.df(phillips_de_lm_all_resid, type = adf_determ,
                           lags = adf_maxlags, selectlags = adf_selcrit)
summary(phillips_de_lm_all_resid.adf)

# ------------------------------------------------------------------------
```

```
# DOLS - Schätzung
# ------------------------------------------------------------------------------
if (!require(dynlm)) install.packages("dynlm")
library(dynlm)
summary(dynlm(INFL_DE ~ AL_DE + L(diff(AL_DE),1:4) + L(diff(AL_DE),-1:-4), data_phillips_de_zoo))

# ------------------------------------------------------------------------------
#                                 Ende
# ------------------------------------------------------------------------------
```

Listing A.1: ./R-code-ss26/ZOE_ss26_vorl_folien_R–1_Phillips.R

A.2. Börsenkurse USA

Verwendung:

- Deskriptiv in 1.1.2,
- Autoregressiv in 1.1.2.

Daten: `ie_data_2026_02_22.xls`, ursprüngliche Quelle: siehe Abbildung 1.2.

R-Programm:

```
################################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 2
################################################################################
#   # Folie 12: Prognose von Aktienindizes
#   Ersteller: RT, 2019_04_25, 2021_04_12, 2022_04_07, 2025_04_25, 2026_02_22
# ------------------------------------------------------------------------------

# ------------------------------------------------------------------------------
# Vorspann
# ------------------------------------------------------------------------------

# Lege Verzeichnis fest, in dem R-Programme stehen oder verwende in R Studio
# den Befehl Set Working Directory -> To Source File Location
#setwd("")


# Logische Variable, die festlegt, ob von Graphiken PDF-Dateien erzeugt werden.
print_pdf = FALSE       # Werte: TRUE, FALSE

# Installiere benötigte R packages und lade diese
    # Package zum Lesen von Excel-Dateien, das kein Java benötigt
if (!require(readxl)) install.packages("readxl")
library(readxl)
```

```
    # Package für spezielle Dataframes für Zeitreihendaten
    # Im Gegensatz zu ts kann zoo auch tägliche Daten handhaben
if (!require(zoo)) install.packages("zoo")
library(zoo)

    # Package für KQ-Schätzung von dynamischen Regressionsmodellen und
    # AR-Modellen
if (!require(dynlm)) install.packages("dynlm")
library(dynlm)

# Lade Funktion, die Modellselektionskriterien wie in EViews berechnet
source("SelectCritEViews.R")

# ------------------------------------------------------------------------
# Lese Daten aus Excel-Datei ein, erstelle zoo Dataframe
# ------------------------------------------------------------------------
    #für Daten bis 2013, um Graphik im Skript, Version 2014 zu reproduzieren
# data_aktien <- data.frame( read_excel(path="Daten/ie_data_2013_04_14.xls",
#                     range="Data!H9:J1812",  #J1812 für ie_data_2021_04_12.xls,
#                                             # J1716 für ie_data_2013_04_14.xls
#                                             # um Graphik im Skript zu erstellen
#                     col_names=c("RealPrice", "RealDividend", "RealEarnings")))

    # Daten bis zum aktuellen Rand
# data_aktien <- data.frame( read_excel(path="Daten/ie_data_2022_04_07.xls",
#                     range="Data!H9:K1824",
#                     col_names=c("RealPrice", "RealDividend",
#                                 "RealTotalReturnPrice", "RealEarnings")))
```

```
data_aktien <- data.frame(
  read_excel(path="Daten/ie_data_2026_02_22.xls",
             range="Data!H9:K1869",
             col_names=c("RealPrice", "RealDividend",
                         "RealTotalReturnPrice", "RealEarnings")))

    # überprüfe Einlesen
head(data_aktien)
tail(data_aktien)

    # Erstelle zoo-Dataframe via ts-Dataframe
data_aktien_zoo <- zoo(ts(data_aktien, start = c(1871, 1), frequency = 12))

    # überprüfe Umwandlung: zeige die ersten 13 Beobachtungen an
data_aktien_zoo[1:13,]
tail(data_aktien_zoo)

# ------------------------------------------------------------------------------
# Erstelle Graphik der Zeitreihen für realen Aktienindex und reale Gewinne
# ------------------------------------------------------------------------------
if (print_pdf) pdf("gr_real_s&p.pdf", 12, 6)
par(mai = c(1.0, .8, 0.8, 0.8)) # Rahmen rechts etwas vergrößern
plot(data_aktien_zoo[,c("RealPrice")], col = "blue",
            ylim = c(0, max(data_aktien_zoo[,c("RealPrice")])), ylab = "",
            xlab = "Year")
grid()

par(new = T)        # Befehl für überlappende Graphik

plot(data_aktien_zoo[,c("RealEarnings")], ylab = "y", yaxt = "n",
```

```
     ylim = c(0, max(data_aktien_zoo[,c("RealPrice")])/10), col = "green",
     ann = F) # ann = F: keine Achsenbeschriftung, da bereits vorhanden

# Beschriftung der Achsen und Legende:

usr <- par("usr") # Abspeicherung der Koordinaten der Ränder für senkrechte Achsenbeschriftung

text(usr[2] + 0.05 * diff(usr[1:2]), mean(usr[3:4]), "Real S&P Composite Earnings",
     srt = 90, xpd = TRUE, col = "green") # xpd = T: Text auch außerhalb des Plots möglich

text(usr[1] - 0.05 * diff(usr[1:2]), mean(usr[3:4]), "Real S&P 500 Stock Price Index",
     srt = 90, xpd = TRUE, col = "blue")

legend("topleft", c("Price (left scale)", "Earnings (right scale)"),
       bty = "n", col = c("blue", "green"), lty = 1)

axis(side=4)
if (print_pdf) dev.off()

# ------------------------------------------------------------------------------
# Berechne KQ-Schätzer für Regressionsgleichung eines AR(1)-Modells
# ------------------------------------------------------------------------------
    # dynlm erfordert package dynlm
    # L funktioniert wie gewohnt - im Gegensatz zu lag()
rprice_dynlm <- dynlm(RealPrice ~ L(RealPrice), data = data_aktien_zoo)
summary(rprice_dynlm)
SelectCritEViews(rprice_dynlm)

# alternativ
ar(data_aktien_zoo$RealPrice, aic=FALSE, order.max=1, demean=TRUE, method="ols")
```

```
# ------------------------------------------------------------------------
#                                  Ende
# ------------------------------------------------------------------------
```

Listing A.2: ./R-code-ss26/ZOE_ss26_vorl_folien_R–2_SP500.R

A.3. Deutsches Bruttonationalprodukt und deutsches reales Bruttoinlandsprodukt

Verwendung:

- Deskriptiv in 1.2 und 3.
- Schätzungen von verschiedenen deterministischen Trends in 3.2.1 bis 3.2.3.

- Schätzung von Modell mit Saisondummies und Trend in 3.4.

Daten:

- `GermanGNP.dat`, ursprüngliche Quelle: siehe Abbildung 1.3.
- `BBNZ1.Q.DE.N.H.0000.L.xlsx`, ursprüngliche Quelle: siehe Abbildung 1.4.

R-Programm:

```
##############################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 3
##############################################################################
#   Folien 15 (und 80), 15, Vierteljahresdaten des deutschen
#   nominalen und realen Bruttonationalprodukts
#   Ersteller: RT, 2019_04_25, basierend auf Patrick Kratzer
#   Änderungen: RT, 2022_04_19, 2025_04_25 (Aktualisierung reale BIP-Daten,
#   Schätzen von linearem, quadratischem und exponentiellem Trend)
#   RT, 2025_05_19 (Korrektur von Einlesen von rBIP,
#       Alternative zu trend() oder time() in dynlm,
#       Modell mit linearem Trend und Saisondummies mit und ohne dynlm)
#       2026_02_22 (Aktualisierung Daten von Bundesbank)
# ----------------------------------------------------------------------------
```

```
# ------------------------------------------------------------------------------
# Vorspann
# ------------------------------------------------------------------------------

# Lege Verzeichnis fest, in dem R-Programme stehen oder verwende in R Studio
# den Befehl Set Working Directory -> To Source File Location
#setwd("")

# Installiere benötigte R packages und lade diese
    # Package zum Lesen von Excel-Dateien, das kein Java benötigt
if (!require(readxl)) install.packages("readxl")
library(readxl)
    # Package für dynamische Regressionsmodelle
if (!require(dynlm)) install.packages("dynlm")
library(dynlm)

# Logische Variable, die festlegt, ob von Graphiken PDF-Dateien erzeugt werden.
print_pdf = TRUE       # Werte: TRUE, FALSE

# ------------------------------------------------------------------------------
# Lese Daten aus Text-Datei und Excel-Datei ein, erstelle ts Dataframe und Plot
# ------------------------------------------------------------------------------

# Folie 16 (und 82): Vierteljahresdaten des deutschen Bruttonationalprodukts
BNP <- ts(read.table("Daten/GermanGNP.dat", skip = 8, header = TRUE),
          start = c(1975, 1), frequency = 4)

if (print_pdf) pdf("gr_BNP_1975.pdf", 6, 6)
plot(BNP, col = "blue", xlab = "Jahr", ylab = "", main = "Bruttonationalprodukt")
```

```
grid()
if (print_pdf) dev.off()

# # Folie 17: Vierteljahresdaten des deutschen realen Bruttoinlandsprodukts 1970-2010
# bbk_JQ5000 <- data.frame(read_excel(path="Daten/bbk_JQ5000.xls",
#                                     range="BIP!A1:B165",
#                                     col_names=TRUE))
# rBIP <- ts(bbk_JQ5000$BIP, start = 1970, frequency = 4)
#
# if (print_pdf) pdf("r_rBIP_1970.pdf", 6, 6)
# plot(rBIP, col = "blue", ylab = "", xlab = "Jahr", main = "reales Bruttoinlandsprodukt", )
# grid()
# if (print_pdf) dev.off()
   # RT 2025_05_19: range korrigiert zu rBIP!10:B229
# Folie 17: Vierteljahresdaten des deutschen realen Bruttoinlandsprodukts 1970-2025
BBK_all <- data.frame(read_xlsx(path="Daten/BBNZ1.Q.DE.N.H.0000.L.xlsx",
                                range="BBNZ1.Q.DE.N.H.0000.L!B10:B233",
                                col_names="rBIP"))
rBIP <- ts(BBK_all$rBIP, start = 1970, frequency = 4)

if (print_pdf) pdf("gr_rBIP_1970-1_2025-4.pdf", 6, 4)
plot(rBIP, main = "Bruttoinlandsprodukt, preisbereinigt", ylab = "BIP, preisbereinigt")
grid()
if (print_pdf) dev.off()

# ------------------------------------------------------------------------
# Berechne KQ-Schätzer für einfaches Regressionsmodell mit linearem Trend
# ------------------------------------------------------------------------
# Vgl. Folie 86
```

```
# Verwendung von dynlm: verschiedene Möglichkeiten # 2025_05_19, RT ergänzt Anfang
# a) time(rBIP)
# b) trend(rBIP)

# a) entspricht (ohne dynlm) - beachte Änderung des Konstante
(trend_y_q <- index(rBIP))
summary(lm(rBIP ~ trend_y_q))
  #               (mit dynlm)
summary(dynlm(rBIP ~ time(rBIP)))

# b) entspricht (ohne dynlm)
trend_q <- (1:length(rBIP))/4
summary(lm(rBIP ~ trend_q))
  #             (mit dynlm) # 2025_05_19, RT, ergänzt Ende
rBIP_lin <- dynlm(rBIP ~ trend(rBIP)) # beachte: Konstante ist automatisch dabei
summary(rBIP_lin)           # KQ-Output

# Erstelle Plot mit Zeitreihe, linearem und quadratischem Trend
plot(rBIP, col = "blue", ylab = "", xlab = "Jahr",
     main = "reales Bruttoinlandsprodukt" )
lines(fitted(rBIP_lin))     # füge dem bestehendem Plot eine Gerade mit
# geschätztem linearem Trend hinzu
"Interpretieren Sie den geschätzten Trendparameter."
"Wie können Sie die Abweichungen vom linearen Trend interpretieren?"
"Welche Konsequenzen ergeben sich hieraus für die Eigenschaften der Residuen?"
# plot(residuals(rBIP_lin), ylab = "Residuen", xlab = "Jahr", main ="reales BIP, linearer Trend")
  # RT 2025_05_19 Plot Residuen hinzugefügt
# ------------------------------------------------------------------------
# Berechne KQ-Schätzer für einfaches Regressionsmodell mit quadratischem Trend
# ------------------------------------------------------------------------
```

```
# Vgl. Folie 89
rBIP_quad <- dynlm( rBIP ~ trend(rBIP) + I(trend(rBIP)^2) )
summary(rBIP_quad)           # KQ-Output
lines(fitted(rBIP_quad), col = "red")
# füge dem bestehendem Plot eine Gerade mit
# geschätztem quadratischem Trend hinzu
legend("bottomright", c("reales BIP", "geschätzter linearer Trend",
                    "geschätzter quadratischer Trend"), lty = 1,
       col = c("blue","black", "red"))

# Erstelle Plot mit Residuen für beide geschätzte Trendarten
par(mfrow = c(1,2))  #Ermögliche zwei Plots in einem Graphikfenster
plot(residuals(rBIP_lin), ylab = "Residuen", main = "linearer Trend")
plot(residuals(rBIP_quad), ylab = "Residuen", main = "quadratischer Trend")



# ------------------------------------------------------------------------
# Berechne KQ-Schätzer für einfaches Regressionsmodell mit exponentiellem Trend
# ------------------------------------------------------------------------
# Vgl. Folie 87
par(mfrow = c(1,2)) # Ermögliche zwei Plots in einem Graphikfenster
plot(log(rBIP), col = "blue", ylab = "", xlab = "Jahr",
     main = "log. reales Bruttoinlandsprodukt" )
# erstelle Plot mit logarithm. Daten
rBIP_exp <- dynlm(log(rBIP) ~ trend(rBIP)) # beachte: Konstante ist automatisch dabei
summary(rBIP_exp)            # KQ-Output
lines(fitted(rBIP_exp))      # füge dem bestehendem Plot eine Gerade mit
# geschätztem exp. Trend hinzu
legend("bottomright", c("log. reales BIP", "lin. Trend"),
       lty = 1, col = c("blue","black"))
```

```
plot(residuals(rBIP_exp), ylab = "Residuen", main = "Residuen")
"Interpretieren Sie den geschätzten Trendparameter."
"Ist der Schätzer des Trendparameters erwartungstreu?"
paste("Dürfen Sie, gegeben den Plot der Residuen,",
      "z.B. auf Signifikanz des Trendparameters testen?")

(exp(coef(rBIP_exp)[2]) -1) * 100 # exakte Interpretation bei log-level-Modell

# ------------------------------------------------------------------------------
# Berechne KQ-Schätzer für einf. Reg.-modell mit linearem Trend und Saisondummies
# ------------------------------------------------------------------------------

# Vgl. Folie 104: Erzeuge Saisondummies
# Referenzquartal:Frühjahr - entspricht bei Verwendung von index(rBIP) 1970.00
    # ohne dynlm
S_D <- ((index(rBIP) - trunc(index(rBIP))) == 0.25)   # Sommerquartal
H_D <- ((index(rBIP) - trunc(index(rBIP))) == 0.50)   # Herbstquartal
W_D <- ((index(rBIP) - trunc(index(rBIP))) == 0.75)   # Winterquartal
summary(lm(rBIP ~ trend_y_q + S_D + H_D + W_D))
    # mit dynlm
rBIP_lin_saison <- dynlm(rBIP ~ time(rBIP) + season(rBIP))
summary(rBIP_lin_saison)
  # Erstelle Plot mit gefitteter Regressionsgerade und Residuenplot
par(mfrow = c(1,2)) # Ermögliche zwei Plots in einem Graphikfenster
plot(rBIP, col = "blue", ylab = "", xlab = "Jahr",
     main = "reales Bruttoinlandsprodukt" )
lines(fitted(rBIP_lin_saison))     # füge dem bestehendem Plot eine Gerade mit
# geschätztem linearem Trend und Saisondummies hinzu
plot(residuals(rBIP_lin_saison), ylab = "Residuen", main = "lin. Trend & Saisondummies")
```

```
# ------------------------------------------------------------------------
#                                  Ende
# ------------------------------------------------------------------------
```

Listing A.3: ./R-code-ss26/ZOE_ss26_vorl_folien_R–3_BIP.R

A.4. Biernachfrage in den Niederlanden

Verwendung:

- Beschreibung des Datensatzes in 2,
- Statisch in 2.3.1,
- FDL-Modell in 2.3.2,
- FDL-Modell nach log-Transformation in 2.4.2.

Daten: `beer_dat.txt`, ursprüngliche Quelle siehe Tabelle 2.1.

R-Programm:

```
##############################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 4
##############################################################################
#   Folien 24-25, 41-45, 52-53 Analyse des Bierabsatzes in Abhängigkeit von
#                           Werbeausgaben
#   Ersteller: RT, 2019_04_25 basierend auf Patrick Kratzer
#              RT, 2021_04_22, 2022_04_19, 2025_05_11 (log (temp + 3))
# ----------------------------------------------------------------------------

# ----------------------------------------------------------------------------
# Vorspann
# ----------------------------------------------------------------------------

# Lege Verzeichnis fest, in dem R-Programme stehen oder verwende in R Studio
# den Befehl Set Working Directory -> To Source File Location
#setwd("")

# Logische Variable, die festlegt, ob von Graphiken PDF-Dateien erzeugt werden.
print_pdf = FALSE       # Werte: TRUE, FALSE

# Installiere benötigte R packages und lade diese
    # Package für KQ-Schätzung von dynamischen Regressionsmodellen und
    # AR-Modellen
if (!require(dynlm)) install.packages("dynlm")
library(dynlm)
```

```
# Lade Funktion, die Modellselektionskriterien wie in EViews berechnet
source("SelectCritEViews.R")

# ------------------------------------------------------------------------
# Lese Daten aus Text-Datei ein, erstelle ts Dataframe
# ------------------------------------------------------------------------

beer     <- read.table("Daten/beer_dat.txt", header = TRUE)
beer_dat <- ts(beer, start = c(1978, 1), freq = 6)

# ------------------------------------------------------------------------
# Berechne KQ-Schätzer für statisches Regressionsmodell, S. 41f
# ------------------------------------------------------------------------

ols_beer_1 <- lm(q ~ wa + temp, data = beer_dat)
summary(ols_beer_1)
SelectCritEViews(ols_beer_1)



if (print_pdf) pdf("gr_resid_beer.pdf", 6, 6)
plot(residuals(ols_beer_1),
     type = "l", col = "blue",
     xlab = "Zeit", ylab = "Residuen")
grid()
abline(h = 0, col = "red", lty = 2)
abline(h =  summary(ols_beer_1)$sigma, lty = 2, col = "red")
abline(h = -summary(ols_beer_1)$sigma, lty = 2, col = "red")
if (print_pdf) dev.off()
```

```
# ------------------------------------------------------------------------------
# Berechne KQ-Schätzer für Finite Distributed Lag Modell, S. 44f
# ------------------------------------------------------------------------------

ols_beer_2 <- dynlm(q ~ 1 + L(wa, 0:2) + temp, data = beer_dat)
summary(ols_beer_2)
SelectCritEViews(ols_beer_2)

if (print_pdf) pdf("gr_resid_beer_fdl.pdf", 6, 6)
plot(residuals(ols_beer_2),
     type = "l", col = "blue",
     xlab = "time", ylab = "residuals")
grid()
abline(h = 0, col = "red", lty = 2)
abline(h =  summary(ols_beer_2)$sigma, lty = 2, col = "red")
abline(h = -summary(ols_beer_2)$sigma, lty = 2, col = "red")
if (print_pdf) dev.off()

# ------------------------------------------------------------------------------
# Berechne KQ-Schätzer für Finite Distributed Lag log-log-Modell, S. 52f
# ------------------------------------------------------------------------------
# abweichend von Folien temp nicht logarithmiert

ols_beer_3a <- dynlm(log(q) ~ L(log(wa),0:2) + log(temp), data = beer_dat)
summary(ols_beer_3a)
SelectCritEViews(ols_beer_3a)

# jetzt auch temp logarithmiert und Beobachtungen mit temp <= 0 eliminiert
beer_dat_zoo = zoo(beer_dat)
```

```
beer_dat_zoo$temp <- beer_dat_zoo$temp + 3   # added 2025_05_11 RT
ols_beer_3b <- dynlm(log(q) ~ L(log(wa),0:2) + log(temp), data = beer_dat_zoo)
summary(ols_beer_3b)
SelectCritEViews(ols_beer_3b)

# ------------------------------------------------------------------------------
# Teste langfristige Elastizität im Finite Distributed Lag log-log-Modell, S. 52f
# ------------------------------------------------------------------------------

library(car)
linearHypothesis(ols_beer_3b, hypothesis.matrix = c(0,1,1,1,0))

    # doing it via rearranging the regression equation
ols_beer_3b_test <- dynlm(log(q) ~ 1 +
                            I(L(log(wa), 0) - L(log(wa), 2)) +
                            I(L(log(wa), 1) - L(log(wa), 2)) +
                            L(log(wa), 2) + log(temp), data = beer_dat_zoo)
summary(ols_beer_3b_test)
SelectCritEViews(ols_beer_3b_test)

# ------------------------------------------------------------------------------
#                                Ende
# ------------------------------------------------------------------------------
```

Listing A.4: ./R-code-ss26/ZOE_ss22_vorl_folien_R–4_beer.R

A.5. Weiterbildung und Ausschussquote

Verwendung:

- Beschreibung des Datensatzes in 2.4.2,
- OLS-Modell in 2.4.2.

Daten: `jtrain` in R-Paket `wooldridge`

R-Programm:

```
##############################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 5
##############################################################################
#   Folien 57, 58 Weiterbildung und Ausschussquote: Datensatz jtrain in
#       R-Paket wooldridge
#   Ersteller: RT, 2022_04_20
# ----------------------------------------------------------------------------

# ----------------------------------------------------------------------------
# Vorspann
```

```
# ------------------------------------------------------------------------

# Lege Verzeichnis fest, in dem R-Programme stehen oder verwende in R Studio
# den Befehl Set Working Directory -> To Source File Location
#setwd("")

# Installiere benötigte R packages und lade diese

# Paket zum Bereitstellen der Wooldridge-Daten
if (!require(wooldridge)){
    install.packages("wooldridge")
}
library("wooldridge")

# ------------------------------------------------------------------------
# Lese Daten aus R-Paket ein
# ------------------------------------------------------------------------

data('jtrain')

# ------------------------------------------------------------------------
# Berechne KQ-Schätzer für Daten für das Jahr 1987
# ------------------------------------------------------------------------

jtrain_subset_1987 <- subset(jtrain, subset = (jtrain$year == 1987))
        # Selektiere Beobachtungen für das Jahr 1987

summary( lm(log(scrap) ~ hrsemp + log(sales) + log(employ),
            data = jtrain_subset_1987) )
```

```
# ------------------------------------------------------------------------
#                                  Ende
# ------------------------------------------------------------------------
```

Listing A.5: ./R-code-ss26/ZOE_ss22_vorl_folien_R–5_jtrain.R

A.6. Regensburger Mietspiegel

Verwendung:

- OLS mit quadratischem Alter in 2.4.3.

Daten für Unterrichtszwecke nicht verfügbar.

A.7. Löhne und individuelle Charakteristika

Verwendung:

- Mit Dummies für Untergruppen in 2.5.2,
- Mit Interaktionstermen in 2.5.4.

Daten: `wage1` in R-Paket `wooldridge`

R-Programm:

```
##############################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 6
##############################################################################
#   Lohnbeispiel mit mehreren Dummies, Folien 69f und 76f
#   Ersteller: RT, 2019_05_09, basierend auf Patrick Kratzer
# ----------------------------------------------------------------------------

# ----------------------------------------------------------------------------
# Vorspann
# ----------------------------------------------------------------------------
```

```
# Lege Verzeichnis fest, in dem R-Programme stehen
#setwd("")
#oder
# verwende in RStudio -> Session -> Set Working Directory -> To Source File Loc.

# Lade Funktion zum Berechnen der Modellselektionskriterien a la EViews
source("SelectCritEViews.R")

# ------------------------------------------------------------------------------
# Lese Daten aus Text-Datei ein und erstelle Dummy-Variablen für Untergruppen
# ------------------------------------------------------------------------------
wage1_data <- read.table("Daten/wage1.txt", header = TRUE)
attach(wage1_data)

femmarr   <- female * married
malesing <-(1-female) * (1- married)
malemarr <-(1-female) * married

# ------------------------------------------------------------------------------
# Berechne KQ-Schätzer mit Dummy-Variablen für Untergruppen
# ------------------------------------------------------------------------------
ols_wage_1 <- lm(log(wage) ~ 1 + femmarr + malesing + malemarr + educ +
                     exper + I(exper^2) + tenure + I(tenure^2))
summary(ols_wage_1)
SelectCritEViews(ols_wage_1)

    # Testen, ob unbedingter Lohnunterschied
    # zwischen Männer und Frauen signifikant? (S. 71)
    # Nullhypothese: beta_0 + beta_1 = 0
```

```
    # Alternative : beta_0 + beta_1 /= 0
library(car)
linearHypothesis(ols_wage_1, "1*malemarr - 1*femmarr = 0")
# ------------------------------------------------------------------------------
# Berechne KQ-Schätzer mit Interaktionstermen, Folie 76f
# ------------------------------------------------------------------------------


ols_wage_2 <- lm(log(wage) ~ 1 + female + educ + exper + I(exper^2) +
                      tenure + I(tenure^2) + I(female*educ))
summary(ols_wage_2)
SelectCritEViews(ols_wage_2)

# ------------------------------------------------------------------------------
#                                 Ende
# ------------------------------------------------------------------------------
```

Listing A.6: ./R-code-ss26/ZOE_ss22_vorl_folien_R–6_Wage_dummies.R

A.8. Regressionen mit trendbehafteten Variablen

Verwendung:

- Fall II: Zusammenhang zwischen Abweichungen in 3.3.2.
- Fall III: Scheinregression in 3.3.3.

R-Programm:

```
################################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 7
################################################################################
#   Programm zur Illustration zu linearen Regressionen mit Variablen,
#   die Trends und/oder Saisonalität aufweisen,
#   siehe Folien Seite 90 folgende.
#
#   Hinweis: Trends können einfach in dem R-Paket dynlm integriert werden.
#   Siehe R-Programm zu Vorlesungsfolien - Nummer 3
#
#   Illustration anhand von simulierten Daten
#   Ersteller: RT, 2021_05_07
# ------------------------------------------------------------------------------

# ------------------------------------------------------------------------------
#   Definiere Parameter
# ------------------------------------------------------------------------------

n       <- 10000        # Stichprobengröße
beta_0  <- 1            # beta_0 und beta_1 für Zusammenhang in Abweichungen,
beta_1  <- 0.8          # siehe Gleichung (3.3) im Skript
sigma_u <- 2            # Standardabweichung für Fehler in (3.3)
```

```
delta_0 <- 0            # Trendparameter für x_t, siehe (3.2)
delta   <- 0.05
sigma_xtilde <- 1       # xtilde_t wird als iid N(0,sigma_xtilde^2) erzeugt
alpha_0 <- 0            # Trendparameter für y_t, siehe (3.1)
alpha   <- 0.05
sigma_ytilde <- 2       # ytilde_t wird als iid N(0,sigma_xtilde^2) erzeugt

# ------------------------------------------------------------------------
#   Generiere trendbehaftete Regressorvariable für Fall II und III
# ------------------------------------------------------------------------

set.seed(42)
#   Erzeuge trendbehaftete x und plotte xtilde und x
xtilde  <- rnorm(n, 0, sigma_xtilde)
x       <- delta_0 + delta * (1:n) + xtilde
plot(xtilde, type = "l")
plot(x, type = "l")

# ------------------------------------------------------------------------
#   Fall II, Seite 93 folgende: Zusammenhang zwischen Abweichungen
# ------------------------------------------------------------------------

#   Generiere ytilde gemäß (3.3)
u       <- rnorm(n, mean = 0, sd = sigma_u)
ytilde  <- beta_0 + beta_1 * xtilde + u

#   Generiere y gemäß (3.1) mit dieser ytilde-Variablen
y       <- alpha_0 + alpha * (1:n) + ytilde

#   Plotte Streudiagramm (Scatterplot)
```

```
plot(x, y)


#   Schätze korrekt spezifizierte einfache lineare Regression gemäß (3.3)
tilde_ols <- lm(ytilde ~ xtilde)
summary(tilde_ols)

#   Schätze fehlspezifizierte einfache lineare Regression in den
#   trendbehafteten Variablen wie auf Seite 91
level_ols <- lm(y ~ x)
summary(level_ols)

#   Schätze korrekt spezifzierte lineare Regression in Niveaus
#   einschließlich Trendterm, siehe Seite 94
level_trend_ols <- lm(y ~ x + I(1:n))
summary(level_trend_ols)



# ------------------------------------------------------------------------------
#   Fall III, Seite 98 folgende: Scheinregression
# ------------------------------------------------------------------------------

#   Erzeuge trendbehaftete y und plotte ytilde und y
ytilde  <- rnorm(n, 0, sigma_ytilde)
y       <- alpha_0 + alpha * (1:n) + ytilde
plot(ytilde, type = "l")
plot(y, type = "l")

#   Plotte Streudiagramm (Scatterplot)
```

```
plot(x, y)

#   Schätze einfache lineare Regression in Niveauvariablen
schein_ols <- lm(y ~ x)
summary(schein_ols)


#   Schätze lineare Regression mit Trend, um Problem der Scheinregression
#   zu vermeiden
schein_trend_ols <- lm(y ~ x + I(1:n))
summary(schein_trend_ols)

# ------------------------------------------------------------------------
#                                Ende
# ------------------------------------------------------------------------
```

Listing A.7: ./R-code-ss26/ZOE_ss22_vorl_folien_R–7_Trends_u_Regression.R

A.9. Monte-Carlo-Simulation eines AR(1)-Prozesses

Verwendung in Abschnitt 4.3.1.

R-Programm:

```
##############################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 8
##############################################################################
#   Programm zur Generierung von R Replikationen eines
#   AR(1)-Prozesses ohne Konstante
#   Folien 131f
#   Ersteller: RT, 2019_05_23, 2023_04_18
# ----------------------------------------------------------------------------

# ----------------------------------------------------------------------------
# Vorspann
# ----------------------------------------------------------------------------

# graphics.off()      # Schließe alle Graphikfenster

# Lege Verzeichnis fest, in dem R-Programme stehen
# setwd("")
# oder
# verwende in RStudio -> Session -> Set Working Directory -> To Source File Loc.

# ----------------------------------------------------------------------------
# Definiere Parameter für Generierung eines AR(1)-Prozesses
# ----------------------------------------------------------------------------

# Logische Variable, die festlegt, ob von Graphiken PDF-Dateien erzeugt werden.
print_pdf = FALSE       # Werte: TRUE, FALSE

# Setze Parameter des Modells für eine Realisation (Folien 137ff)

set.seed(42)        # Randomseed
```

```
T     <- 50          # Stichprobengröße

nu    <- 0           # Konstante (impliziert Mittelwert von mu = nu (1-alpha)
alpha <- 0.9         # Parametervektor
sigma <- 1           # Standardabweichung des Fehlers
y0    <- 0           # Startwert des AR(1)-Prozesses


# ------------------------------------------------------------------------------
# Generiere eine Realisation eines AR(1), Folie 130f
# ------------------------------------------------------------------------------

# Generieren einer Realisation eines AR(1)-Prozesses
e <- rnorm(T+1, mean = 0, sd = sigma)   # Ziehen von u
y <- rep(y0, T+1)    # Initialisiere Vektor mit Zeitreihenbeobachtungen
                     # mit Startwert y_0; RT 2023_04 simplified

# Erzeugen der y_t-Beobachtungen
  # Möglichkeit A) Schleife zur Generierung der Beobachtungen

for (t in (2:(T+1))){
    y[t] <- nu + y[t-1] * alpha + e[t]
}
y <- y[-1]  # Eliminiere Startwert

  # Möglichkeit B) Verwenden des filter-Befehls in R
  #(schneller, aber auskommentiert)
# y <- filter(nu + e[2:(T+1)], filter = alpha,
#                 method = "recursive", init = y0)
```

```
# Erzeuge Zeitreihe der Realisation
if (print_pdf) pdf("gr_realisation_AR1.pdf", 6, 6)
  plot(y, xlab = "Perioden", ylab=expression(y[t]),
       main =  "Realisation eines AR(1)-Prozesses",
       type="l")
if (print_pdf) dev.off()

# ------------------------------------------------------------------------------
#                                Ende
# ------------------------------------------------------------------------------
```

Listing A.8: ./R-code-ss26/ZOE_ss22_vorl_folien_R–8_AR1_MC.R

A.10. Monte-Carlo-Simulation von vielen AR(1)-Prozessen

Verwendung in Abschnitt 4.3.2.

R-Programm:

```
##############################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 9
##############################################################################
#   Generiert R Realisationen eines AR(1) Prozesses ohne und mit Konstante
```

```
#   und liefert jeweils eine Graphik inklusive unbedingtem Erwartungswert.
#   Folien: 133f
#   Ersteller: RT, 2022_04_22
# ------------------------------------------------------------------------------

# ------------------------------------------------------------------------------
# Vorspann
# ------------------------------------------------------------------------------

# Lege Verzeichnis fest, in dem R-Programme stehen
# setwd("")
# oder
# verwende in RStudio -> Session -> Set Working Directory -> To Source File Loc.

# Logische Variable, die festlegt, ob von Graphiken PDF-Dateien erzeugt werden.
print_pdf = FALSE             # Werte: TRUE, FALSE

# ------------------------------------------------------------------------------
#           Definiere Parameter für AR(1)-Prozess ohne Konstante
# ------------------------------------------------------------------------------

nu              <- 1        # Konstante
alpha           <- 0.9      # AR-Parameter
sigma           <- 1        # Standardabweichung Fehlerterm
y_0             <- 5        # Startwert

T               <- 101      # Stichprobengröße
R               <- 50       # Zahl der Replikationen

set.seed(12345)             # Seed Value
```

```
# ------------------------------------------------------------------------
#          Generiere R mal AR(1)-Prozess ohne Konstante
# ------------------------------------------------------------------------

if (print_pdf) pdf("gr_ar1_R_50.pdf", 12, 6)
par(mfrow = c(1,2))  # Erstelle zwei Graphiken in einer Datei

        # Erzeuge Matrix mit Fehler, wobei in jeder Spalte
        # die Fehler für eine Realisation stehen
e <- matrix(rnorm((T+1) * R), nrow = T+1, ncol = R)
        # entsprechend eine Matrix für die Zeitreihe, wobei
        # alle Werte mit dem Startwert initialisiert werden
y <- matrix(y_0, nrow = T+1, ncol = R)

        # Erzeuge R Realisationen gleichzeitig mit Schleife
for(t in 2:(T+1)){
    y[t,] <- alpha * y[t-1,] + e[t,]}

# Unbedingter Erwartungswert
        # Erstelle Vektor mit unbedingtem Erwartungswert
        # für jede Periode.
mu <- rep(y_0, T+1)
        # Berechne unbedingte Erwartungswerte
for(t in 2:(T+1)) {mu[t] <- alpha * mu[t-1]}

# Erwartungswert, falls Prozess stationär
EW_stat <- 0 / (1 - alpha)

# Grenzen des 95%-Konfidenzintervalls
```

```
konf_o <- mu + sqrt(c(0, cumsum(alpha^((0:(T-1))*2)) )) * sigma * qnorm(0.975)
konf_u <- mu - sqrt(c(0, cumsum(alpha^((0:(T-1))*2)) )) * sigma * qnorm(0.975)

plot(0:(T), y[,1], type = "l", ylim = c(min(y), max(y)),
     xlab = "t", ylab = expression(y[t]),
     main = paste(R, " Realisationen eines AR(1)-Prozesses ohne Konstante"),
     col = rainbow(R)[1])

lines(0:T, mu, col = "blue", lwd = 2, lty = 1)
lines(0:T, konf_u, col = "red", lwd = 2, lty = 2)
lines(0:T, konf_o, col = "red", lwd = 2, lty = 2)
abline(h = EW_stat, col = "green", lwd = 2, lty = 2)

for(j in 2:R) {lines(0:T, y[,j], col = rainbow(R)[j])}

legend("bottomright", c("unbed. Erwartungswert", "95% - Konfidenzintervall", "Erwartungswert bei Stationarität
   "),
       lty = 2, col = c("blue", "red", "green"), bty = "n")

# ------------------------------------------------------------------------
#          Generiere R mal AR(1)-Prozess mit Konstante
# ------------------------------------------------------------------------

    # verwende Fehler von oben

    # erstelle eine Matrix für die Zeitreihe, wobei
    # alle Werte mit dem Startwert initialisiert werden
y_c <- matrix(y_0, nrow = T+1, ncol = R)

# Erzeuge R Realisationen gleichzeitig mit Schleife
```

```
for(t in 2:(T+1)){
    y_c[t,] <- nu + alpha * y_c[t-1,] + e[t,]}

# Unbedingter Erwartungswert
    # Erstelle Vektor mit unbedingtem Erwartungswert
    # für jede Periode.
mu_c <- rep(y_0, T+1)
    # Berechne unbedingte Erwartungswerte
for(t in 2:(T+1)) {mu_c[t] <- nu + alpha * mu_c[t-1]}

#  Erwartungswert, falls Prozess stationär
EW_c_stat <- nu / (1 - alpha)

# Grenzen des 95%-Konfidenzintervalls
konf_c_o <- mu_c + sqrt(c(0, cumsum(alpha^((0:(T-1))*2)) )) * sigma * qnorm(0.975)
konf_c_u <- mu_c - sqrt(c(0, cumsum(alpha^((0:(T-1))*2)) )) * sigma * qnorm(0.975)

plot(0:(T), y_c[,1], type = "l", ylim = c(min(y_c), max(y_c)),
     xlab = "t", ylab = expression(y[t]),
     main = paste(R, " Realisationen eines AR(1)-Prozesses mit Konstante"),
     col = rainbow(R)[1])

lines(0:T, mu_c, col = "blue", lwd = 2, lty = 1)
lines(0:T, konf_c_u, col = "red", lwd = 2, lty = 2)
lines(0:T, konf_c_o, col = "red", lwd = 2, lty = 2)
abline(h = EW_c_stat, col = "green", lwd = 2, lty = 2)

for(j in 2:R) {lines(0:T, y_c[,j], col = rainbow(R)[j])}

legend("bottomright", c("unbed. Erwartungswert", "95% - Konfidenzintervall", "Erwartungswert bei Stationarität
```

```
    "),
        lty = 2, col = c("blue", "red", "green"), bty = "n")

if (print_pdf) dev.off()

# ------------------------------------------------------------------------------
#                                 Ende
# ------------------------------------------------------------------------------
```

Listing A.9: ./R-code-ss26/ZOE_ss22_vorl_folien_R–9_AR1_MC_R_50.R

A.11. Monte-Carlo-Simulationen der Schätzeigenschaften eines AR(1)-Modells

Verwendung in Abschnitt 4.6.

R-Programm:

```
##############################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 10
##############################################################################
#   Programm zur Generierung von R Replikationen eines
```

```
#   AR(1)-Prozesses ohne Konstante und Schätzung des AR-Parameters
#   Folien 143ff
#   Ersteller: RT, 2019_05_23
# ------------------------------------------------------------------------------

# ------------------------------------------------------------------------------
# Vorspann
# ------------------------------------------------------------------------------

# graphics.off()      # Schließe alle Graphikfenster

# Lege Verzeichnis fest, in dem R-Programme stehen
# setwd("")
# oder
# verwende in RStudio -> Session -> Set Working Directory -> To Source File Loc.

# ------------------------------------------------------------------------------
# Definiere Parameter für Generierung eines AR(1)-Prozesses und Anzahl d. Repl.
# ------------------------------------------------------------------------------

# Logische Variable, die festlegt, ob von Graphiken PDF-Dateien erzeugt werden.
print_pdf = FALSE       # Werte: TRUE, FALSE

# Setze Parameter des Modells

set.seed(42)        # Randomseed
T_init<- 50         # Anzahl Beobachtungen für Einschwingphase

nu    <- 0          # Konstante (impliziert Mittelwert von mu = nu (1-alpha)
alpha <- 0.9        # Parametervektor
```

```
sigma <- 1          # Standardabweichung des Fehlers
y0    <- 0          # Startwert des AR(1)-Prozesses

# Parameter für Monte-Carlo-Simulation

R     <- 10000      # Zahl der Replikationen
T_vec <- c(20, 50, 100, 500, 1000, 10000)

if (!require(xtable)){
  install.packages("xtable") # benötigt ab Folie 290
}
library(xtable)

# ------------------------------------------------------------------------
# Initialisiere Outputmatrix, führe Simulationen durch, erstelle Histogramme für
# Folien 143ff
# ------------------------------------------------------------------------

alpha_hat_store <- matrix(NA, nrow = R, ncol = length(T_vec)) # Initialisierung
          # der Matrix mit den simulierten Mittelwerten

if (print_pdf) pdf("gr_ar1_mcarlo.pdf", 6, 4)
par(mfrow = c(2, 3))
for(j in (1:length(T_vec)))
{
  for(i in 1:R)
  {
    T     <- T_vec[j]
    u     <- rnorm(T + T_init, mean = 0, sd = sigma)   # Ziehen von u
    # Generieren der Zeitreihe
```

```
    y     <- filter(nu + u[2:(T + T_init)], filter = alpha,
                    method = "recursive")[T_init+(1:T)]
    # Schätzen von alpha
    alpha_hat_store[i,j] <- coefficients( lm(y[2:T] ~ -1 + y[1:(T-1)]) )[1]

  }
  # Erstellen von Histogrammen
  hist(alpha_hat_store[,j], breaks = 30, freq = F, xlab = "",
       col = "lightblue", main = paste("T = ", T_vec[j]))

}
if (print_pdf) dev.off()

# Erstelle Tabelle mit Mittelwerten und Standardabweichungen
fx <- function(x) {c(mean(x), sd(x))}
table_output <- apply(alpha_hat_store, 2, fx)
# gebe Spalten und Zeilen Namen
rownames(table_output) <- c("Mittelwerte", "Standardabweichungen")
colnames(table_output) <- paste0("T = ", T_vec)
# erstelle Latex-Code für Tabelle
xtable(t(table_output), digits=6)
# Lösche Matrix means aus Simulation
# rm(alpha_hat_store)


# ------------------------------------------------------------------------
# Generiere R Realisationen / Replikationen für einen DGP
# ------------------------------------------------------------------------
# nicht in Folien verwendet.
# Vergleich der Schätzeigenschaften eines Modells ohne Konstante
```

```
# und mit Konstante (überspezifiziert).
# Initialisiere Matrizen zum Speichern der Simulationsergebnisse
    # In einer Zeile stehen die Ergebnisse für eine Replikation (=Stichprobe)
    # Spalte 1: hat(Parameter), Spalte 2: Standardfehler des Parameterschätzers
    # Spalte 3: t-Statistik, Spalte 4: p-Wert
set.seed(42)        # Randomseed
T <- 50

T_all <- T + T_init
    # Für AR(1)-Modell ohne Konstante
alpha_I_hat_store   <- matrix(0, nrow = R, ncol = 4)
    # Für AR(1)-Modell mit Konstante
alpha0_II_hat_store  <- matrix(0, nrow = R, ncol = 4)
alpha1_II_hat_store  <- matrix(0, nrow = R, ncol = 4)

# compute mean of series to generate
mu <- nu / (1 - alpha)

# Schleife über Anzahl der Replikationen
for (r in (1:R))
{
    # Generieren einer Realisation eines AR(1)-Prozesses
    u <- rnorm(T_all, mean = 0, sd = sigma)   # Ziehen von u

    y_all <- filter(nu + u[2:T_all], filter = alpha,
                    method = "recursive")

    y <- y_all[(T_init+1):T_all]
    # Berechnen der KQ-Schätzer
        # Für AR(1)-Modell ohne Konstante
```

```
    ols_I  <- lm(y[2:T] ~ -1 + y[1:(T-1)])
            # Beachte x=y_{t-1]. Deshalb y_t von t=2,...,N
        # Für AR(1)-Modell mit Konstante
    ols_II <- lm(y[2:T] ~ 1 + y[1:(T-1)])

    # Speichern der Parameterschätzungen
        # Für AR(1)-Modell ohne Konstante
    alpha_I_hat_store[r,] <- summary(ols_I)$coefficients[1:2]
        # Für AR(1)-Modell mit Konstante
    alpha0_II_hat_store[r,] <- summary(ols_II)$coefficients[1,]
    alpha1_II_hat_store[r,] <- summary(ols_II)$coefficients[2,]
}

# Berechnen der Mittelwerte der Parameterschätzungen

(colMeans(alpha_I_hat_store))
(colMeans(alpha0_II_hat_store))
(colMeans(alpha1_II_hat_store))

# Erstellen von Histogrammen
hist(alpha_I_hat_store[,1], breaks=sqrt(R),
       xlab = expression(hat(alpha)[1]), main=paste("Histogramm für T= ", T,
                                                    sep=""))
hist(alpha0_II_hat_store[,1],breaks=sqrt(R))
hist(alpha1_II_hat_store[,1],breaks=sqrt(R))

# Erstellen von geschätzten Dichten der Parameterschätzer
plot(density(alpha_I_hat_store[,1]), main="Monte-Carlo-Dichte von hat alpha",
     col="blue")
lines(density(alpha1_II_hat_store[,1]), col="red")
```

```
legend("topleft", c("AR(1)-Modell ohne Konstante",
                    "AR(1)-Modell mit Konstante"), col=c("blue","red"))
# ------------------------------------------------------------------------
#                                Ende
# ------------------------------------------------------------------------
```

Listing A.10: ./R-code-ss26/ZOE_ss22_vorl_folien_R–10_AR1_Est_MC_R_10000.R

A.12. Monte-Carlo-Simulation von Randon Walks

Verwendung:

- Random Walk ohne Drift in Abschnitt 6.2.2,
- Random Walk mit Drift in Abschnitt 6.3.1.

R-Programm:

```
##############################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 11
##############################################################################
```

```
#   Generiert R Realisationen von Random Walks Prozessen ohne und mit Drift
#   und erstellt jeweils eine Graphik
#   Folien: 198ff, 204ff
#   Ersteller: RT, 2022_04_22, 2025_06_23 (mu_1 in AR(1)-Gleichung)
# ------------------------------------------------------------------------------

# ------------------------------------------------------------------------------
# Vorspann
# ------------------------------------------------------------------------------

# Lege Verzeichnis fest, in dem R-Programme stehen
# setwd("")
# oder
# verwende in RStudio -> Session -> Set Working Directory -> To Source File Loc.

# Logische Variable, die festlegt, ob von Graphiken PDF-Dateien erzeugt werden.
print_pdf = FALSE            # Werte: TRUE, FALSE

# ------------------------------------------------------------------------------
#           Definiere Parameter für AR(1)-Prozess ohne Konstante
# ------------------------------------------------------------------------------

mu_1            <- 0.5      # Konstante (für Random Walk mit Drift)
alpha           <- 1        # AR-Parameter
sigma           <- 1        # Standardabweichung Fehlerterm
y_0             <- 0        # Startwert

T               <- 200      # Stichprobengröße
R               <- 10       # Zahl der Replikationen
```

```
set.seed(12345)             # Seed Value

# ------------------------------------------------------------------------
#          Generiere R mal RW-Prozess ohne Drift
# ------------------------------------------------------------------------

        # Erzeuge Matrix mit Fehler, wobei in jeder Spalte
        # die Fehler für eine Realisation stehen
e <- matrix(rnorm((T+1) * R), nrow = T+1, ncol = R)
        # entsprechend eine Matrix für die Zeitreihe, wobei
        # alle Werte mit dem Startwert initialisiert werden
y <- matrix(y_0, nrow = T+1, ncol = R)

        # Erzeuge R Realisationen gleichzeitig mit Schleife
for(t in 2:(T+1)){
    y[t,] <- mu_1 + alpha * y[t-1,] + e[t,]}


if (print_pdf) pdf("gr_rw_R_10.pdf", 12, 6)
plot(0:(T), y[,1], type = "l", ylim = c(min(y), max(y)),
     xlab = "t", ylab = expression(y[t]),
     main = paste(R, " Realisationen eines Random Walks"),
     col = rainbow(R)[1])

for(j in 2:R) {lines(0:T, y[,j], col = rainbow(R)[j])}
if (print_pdf) dev.off()

# ------------------------------------------------------------------------
#          Generiere R mal RW-Prozess mit Drift
# ------------------------------------------------------------------------
```

```
    # erstelle eine Matrix für die Zeitreihe, wobei
    # alle Werte mit dem Startwert initialisiert werden
y_c <- matrix(y_0, nrow = T+1, ncol = R)

    # Erzeuge R Realisationen gleichzeitig mit Schleife
for(t in 2:(T+1)){
    y_c[t,] <- nu + alpha * y_c[t-1,] + e[t,]}

if (print_pdf) pdf("gr_rw_drift_R_10.pdf", 12, 6)
plot(0:(T), y_c[,1], type = "l", ylim = c(min(y_c), max(y_c)),
     xlab = "t", ylab = expression(y[t]),
     main = paste(R, " Realisationen eines Random Walks mit Drift"),
     col = rainbow(R)[1])

for(j in 2:R) {lines(0:T, y_c[,j], col = rainbow(R)[j])}
if (print_pdf) dev.off()

# ------------------------------------------------------------------------
#                                Ende
# ------------------------------------------------------------------------
```

Listing A.11: ./R-code-ss26/ZOE_ss22_vorl_folien_R–11_RW_MC_R_10.R

A.13. Keynesianische Konsumfunktion

- Beispiel in Abschnitt 6.4.3.

A.14. Illustration eines t-Tests auf Autokorrelation bei strenger Exogenität

Verwendung in Abschnitt 8.1.4

R-Programm:

```
##############################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 12
##############################################################################
#
# R-Programm zur Illustration eines einfachen asymptotischen t-Tests auf
# Autokorrelation in den Fehlern bei streng exogenen Regressoren in den Folien
# zu Zeitreihenökonometrie, Seiten 244-246.
```

```
# Hinweise:
# Wenn der dynlm-Befehl verwendet wird, müssen die Daten
# entweder als ein ts- oder ein zoo-Dataframe vorliegen.
#
# Statt dem dynlm-Befehl lässt sich auch der lm-Befehl zusammen
# mit dem embed-Befehl nutzen. Siehe Zeile 84 und folgende.
#
# Die generierten Fehler u_t zentriert, siehe Zeilen 45, 46,
# damit Voraussetzung E[u_t] = 0 in Realisation erfüllt wird.
#
# RT, 2021_07_08, 2022_04_22
#

# ------------------------------------------------------------------------------
# Definiere Parameter des DGP und der Stichprobe
# ------------------------------------------------------------------------------

T       <- 100          # Stichprobengröße
Var_e   <- 1            # Varianz des iid-Fehlers
beta_1  <- 1            # Parameter für lineare Regressionsmodell
rho     <- 0.9          # AR-Koeffizient für Autokorrelation in Fehlern

# ------------------------------------------------------------------------------
# Generiere Stichprobe und berechne KQ-Schätzer
# ------------------------------------------------------------------------------
    # Generieren von i.i.d. Fehlern e_t
    # und autokorrelierten u_t
set.seed(141414)
e       <- rnorm(T, sd = sqrt(Var_e))
u       <- rep(e[1] * sqrt(Var_e / (1-rho^2)), T) #e[1] * sqrt(Var_e / (1-rho^2))
```

```
for (t in (2:T)){
    u[t] <- rho * u[t-1] + e[t]
}
"Mittelwert der generierten u_t's"
mean(u)
    # Zentriere generierte Fehler, damit E[u] = 0 erfüllt ist.
u       <- u - mean(u)
plot(u, type = "l", main = "Autokorrelierte Fehler in DGP verwendet",
     xlab = "Zeit", ylab = expression(u[t]))

    # Generieren von streng exogenen x
    # Verwendung einer Gleichverteilung
x       <- runif(T, min = 2, max = 4)
plot(x, type = "l", main = "Unabhängige Variable in DGP verwendet", xlab = "Zeit",
     ylab = expression(x[t]))

    # Generieren von y mit autokorrelierten Fehlern
y       <- beta_1 * x + u
plot(x, y, main = "Scatterplot", xlab = expression(x[t]), ylab = expression(y[t]))

    # KQ-Schätzung von Regressionsmodell
ols     <- lm(y ~ -1 + x)
summary(ols)

    # Erstelle Vektor mit Residuen aus der KQ-Schätzung
ols_res <- residuals(ols)
    # Plotte Residuen
plot(ols_res, type = "l", xlab = "Zeit", ylab = "Residuen")
# lines(u, col = "red") # ermöglicht Einzeichnen der Fehler des DGP
```

```
# ------------------------------------------------------------------------
# Führe asymptotischen t-Test durch
# ------------------------------------------------------------------------

# Verwende dynlm-Paket. Dafür muss ein ts-Dataframe von den Residuen für dynlm
# erstellt werden.
library(dynlm)
ols_res_ts <- ts(ols_res)
    # t-Statistik wird in Output angegeben
summary(dynlm( ols_res_ts ~ L(ols_res_ts, 1) ))

# Bei starker Autokorrelation schlägt Wooldridge vor, das Modell in
# ersten Differenzen zu schätzen
summary(dynlm(diff(y) ~ diff(x)))

# Alternative Berechnung ohne dynlm-Befehl
# Erstelle Matrix mit (n-1) Zeilen mit y_t in erster Spalte und y_{t-1}
# in zweiter Spalte und verwende lm-Befehl für Autokorrelationstest
u_u_1     <- embed(ols_res, 2)
summary(lm(u_u_1[,1] ~ u_u_1[,2]))

# ------------------------------------------------------------------------
#                                Ende
# ------------------------------------------------------------------------
```

Listing A.12: ./R-code-ss26/ZOE_ss22_vorl_folien_R–12_LM_AR–test.R

A.15. Berechnen eines FLGS-Schätzer und eines Newey-West-Schätzers bei autokorrelierten Fehlern

Verwendung:

- FLGS-Schätzung in Abschnitt 8.4,
- HAC-Schätzung in Abschnitt 8.5.2.

R-Programm:

```
##############################################################################
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 13
##############################################################################
#
#   R-Programm zu Kapitel 8  FGLS und HAC-Standardfehlern für KQ-Schätzung
#   Es erzeugt Stichprobe wie in
#   ZOE_ss22_vorl_folien_R-12_LM_AR-test.R
#   und berechnet dafür sowohl FGLS-Schätzer als auch
#   HAC-Standardfehler für KQ-Schätzer, wobei der Newey-West-Schätzer
#   benutzt wird.
#   RT, 2021_07_16, 2022_04_22
```

```
# ------------------------------------------------------------------------
# Vorspann
# ------------------------------------------------------------------------

# dev.off()             # schließe ein Graphikfenster, falls noch geöffnet
# rm(list=ls())         # lösche Workspace

library(dynlm)          # lädt automatisch auch package zoo und sandwich
# library(sandwich)     # wird für Befehl NeweyWest für HAC-Standardfehler gebraucht
library(lmtest)         # wird für Befehl coeftest für HAC-Standardfehler gebraucht

# ------------------------------------------------------------------------
# Definiere Parameter des DGP und der Stichprobe
# ------------------------------------------------------------------------

####   erzeuge DGP wie in ZOE_vor_folien_R-7-5_S_243_244_LM_AR-test
T       <- 100          # Stichprobengröße
Var_e   <- 1            # Varianz der iid-Fehler
beta_1  <- 1
rho     <- 0.9          # AR-Koeffizient für Autokorrelation in Fehlern

# ------------------------------------------------------------------------
# Generiere Stichprobe
# ------------------------------------------------------------------------

# Generieren von i.i.d. Fehlern e_t
# und autokorrelierten u_t
set.seed(141414)
e       <- rnorm(T, sd = sqrt(Var_e))
u       <- rep(e[1] * sqrt(Var_e / (1-rho^2)), T)
```

```
for (t in (2:T)){
    u[t] <- rho * u[t-1] + e[t]
}
"Mittelwert der generierten u_t's"
mean(u)
# Zentriere generierte Fehler, damit E[u] = 0 erfüllt ist (fehlte in Vorlesung).
u       <- u - mean(u)
plot(u, type = "l", main = "Autokorrelierte Fehler in DGP verwendet",
     xlab = "Zeit", ylab = expression(u[t]))

# Generieren von streng exogenen x
# Verwendung einer Gleichverteilung
x       <- runif(T, min = 2, max = 4)
plot(x, type = "l", main = "Unabhängige Variable in DGP verwendet",
     xlab = "Zeit", ylab = expression(x[t]))

# Generieren von y mit autokorrelierten Fehlern
y       <- beta_1 * x + u
plot(x, y, main = "Scatterplot", xlab = expression(x[t]), ylab = expression(y[t]))

####        Ende wie in ZOE_ss22_vorl_folien_R-12_LM_AR-test.R
####        im Folgenden bei KQ-Schätzung auch Berücksichtigung von Konstante

# ------------------------------------------------------------------------
#   Berechne FGLS-Schätzer
# ------------------------------------------------------------------------
data_zoo <- zoo(cbind(y,x))         # erstelle zoo-Dataframe für die Verwendung von dynlm

#### FGLS-Schätzung
# 1. Schritt: Schätze rho
```

```
mod_dynlm       <- dynlm(y ~ x, data = data_zoo) # Beachte: zoo-dataframe muss explizit
                                                 # übergeben werden, damit die Residuenzeitreihe auch
                                                 # ein zoo-Objekt ist, sonst werden die Vektoren
                                                 # y und x verwendet und somit die Residuen in
                                                 # einem Vektor gespeichert.
summary(mod_dynlm)
resids      <- residuals(mod_dynlm) # Speichere Residuen aus KQ-Schätzung
plot(resids, type = "l", xlab = "Zeit", ylab = expression(hat(u[t])))
    # Schätze Autokorrelation auf Basis der Residuen
mod_res_dynlm   <- dynlm(resids ~ L(resids))
summary(mod_res_dynlm)
rho_hat         <- coefficients(mod_res_dynlm)[2] # Speichere KQ-Schätzer von rho

# 2. Schritt
    # Erzeuge P-Matrix
P_m             <- diag(1, nrow = T)
P_m[1,1]        <- sqrt(1 - rho_hat^2)
for (t in 2:T){
    P_m[t,t-1]  <- - rho_hat
}
    # Berechne "*"-Modell
y_star          <- P_m %*% data_zoo$y
X_star          <- P_m %*% data_zoo$x

    # KQ-Schätzer mit "*"-Modell
mod_fgls        <- lm(y_star ~ X_star)
summary(mod_fgls)
plot(residuals(mod_fgls), type = "l", xlab = "Zeit", ylab = expression(hat(e[t])))

# ------------------------------------------------------------------------
```

```
#   Regressionsoutput mit HAC-Standardfehlern
# ------------------------------------------------------------------------------

# Zu Wahlmöglichkeiten für die Schätzung der heteroskedastischen
# Varianz-Kovarianzmatrix siehe Abschnitt 14.3
library(lmtest)
library(AER)
library(sandwich)
(coeftest(mod_dynlm, vcov = NeweyWest(mod_dynlm)))

# ------------------------------------------------------------------------------
#                                 Ende
# ------------------------------------------------------------------------------
```

Listing A.13: ./R-code-ss26/ZOE_ss22_vorl_folien_R–13_FGLS_HAC.R

A.16. Berechnen von Augmented-Dickey-Fuller-Tests

Verwendung in Abschnitt 9.2

R-Programm:

```
##############################################################################
```

```
#   R-Programm zu Vorlesungsfolien Zeitreihenökonometrie - Nummer 14
##############################################################################
#   berechnet Augmented-Dickey-Fuller-Tests für S&P 500 Daten, die bereits
#   in R-Programm - Nummer 2 (ZOE_ss26_vorl_folien_R-2_SP500.R)verwendet wurden.
#
#   # Folie 286: ADF-Tests von Aktienindizes
#   Ersteller: RT, 2019_07_25, 2022_04_22, 2026_02_22
#
# ----------------------------------------------------------------------------

# ----------------------------------------------------------------------------
# Vorspann
# ----------------------------------------------------------------------------

# Lege Verzeichnis fest, in dem R-Programme stehen oder verwende in R Studio
# den Befehl Set Working Directory -> To Source File Location
#setwd("")

# Installiere benötigte R packages und lade diese
    # Package zum Lesen von Excel-Dateien, das kein Java benötigt
if (!require(readxl)) install.packages("readxl")
library(readxl)

    # Package für spezielle Dataframes für Zeitreihendaten
    # Im Gegensatz zu ts kann zoo auch tägliche Daten handhaben
if (!require(zoo)) install.packages("zoo")
library(zoo)

    # Package für KQ-Schätzung von dynamischen Regressionsmodellen und
    # AR-Modellen
```

```
if (!require(dynlm)) install.packages("dynlm")
library(dynlm)

    # Weitere Pakete zum Ausführen von ADF-Tests werden weiter unten
    # geladen

# ------------------------------------------------------------------------
#   Definiere Parameter für ADF-Tests
# ------------------------------------------------------------------------
    # im Befehl in ur.df (R-Paket urca) verwendet:
adf_determ    <- "trend"    # Auswahl:
                            # "none": keine Konstante, kein Trend,
                            # "drift": nur Konstante, kein Trend
                                # (Fall A in Folien, S. 280),
                            # "trend": Konstante und Trend
                                # (Fall B in Folien, S. 281)
adf_maxlags   <- 10         # maximale Anzahl an Lags in Niveaus
                            # bei automatischer Wahl der
                            # Lagordnung beim ADF-Testen
adf_selcrit   <- "AIC"      # Lagauswahlkriterium: "AIC", "BIC"
                            # falls "Fixed" bestimmt adf_maxlags die AR-Ordnung
                            # der Testgleichung
p             <- 5          # AR-Ordnung
q             <- 0          # MA-Ordnung - currently only 0 allowed

# ------------------------------------------------------------------------
# Lese Daten aus Excel-Datei ein, erstelle zoo Dataframe
# ------------------------------------------------------------------------

data_aktien <- data.frame(
```

```
  read_excel(path="Daten/ie_data_2026_02_22.xls",
             range="Data!H9:K1869",
             col_names=c("RealPrice", "RealDividend",
                         "RealTotalReturnPrice", "RealEarnings")))
head(data_aktien)
tail(data_aktien)

# Erstelle zoo-Dataframe via ts-Dataframe
data_aktien_zoo <- zoo(ts(data_aktien, start = c(1871, 1), frequency = 12))

# überprüfe Umwandlung: zeige die ersten 13 Beobachtungen an
data_aktien_zoo[1:13,]

# Logarithmieren der Daten für "RealPrice"
data_aktien_zoo_select <- log(data_aktien_zoo[,c("RealPrice"), drop = FALSE])

head(data_aktien_zoo_select)

# ------------------------------------------------------------------------------
# Erstelle Graphik der Zeitreihen für realen Aktienindex und reale Gewinne
# ------------------------------------------------------------------------------

plot.zoo(data_aktien_zoo_select, xlab = "Zeit",
         ylab = paste("Logarithmus von ", names(data_aktien_zoo_select)) )

# ------------------------------------------------------------------------------
#           Einheitswurzeltests / Unit Root Tests
# ------------------------------------------------------------------------------

# ADF-Test (Augmented-Dickey-Fuller-Test)
```

```
#   1) verwende R-Paket urca
if (!require(urca)) install.packages("urca")
library(urca)
    # Beachte, dass für alle Schätzungen nur n-lags (=lag.max) Beobachtungen
    # verwendet werden. Deshalb führt eine Änderung der Lagzahl auch zu
    # Änderungen der Teststatistik aufgrund der sich ändernden
    # Zahl an Freiheitsgraden.
    # Im Output gibt tau3 die kritischen Werte für den ADF-test an.
    # Diese hängen von der Stichprobengröße ab und können bei kleinen
    # Stichproben deshalb von den asymptotischen kritischen Werten
    # z.B. Tabelle 9.1 in den Folien abweichen.
data_aktien.adf     <- ur.df(data_aktien_zoo_select, type = adf_determ,
                             lags = adf_maxlags, selectlags = adf_selcrit)
summary(data_aktien.adf)
    # Anzahl der geschätzten Parameter in der Testgleichung
length(data_aktien.adf@testreg$aliased)


#   2) verwende R-Paket tseries
    # gibt auch p-Werte aus, aber nur verfügbar für Spezifikation
    # Konstante und Trend (Fall B in Folien, S. 281)
    # Alternativhypothesen:
    #   - "stationary" entspricht alpha < 1,
    #   - "explosive" entspricht alpha > 1 (nicht in Vorlesung behandelt)
    # Beachte: Die Werte der Teststatistiken können vom Output von ur.df()
    # abweichen, da die verwendete Anzahl an Beobachtungen beim Befehl adf.test()
    # von der Zahl in ur.df() abweichen kann, wenn in ur.df() eine automatische
    # Lagwahl durchgeführt wird.

if (!require(tseries)) install.packages("tseries")
```

```
library(tseries)
if (adf_determ=="trend"){
    # Berechte p-Werte mit adf.test()-Befehl
    print("Note that test statistic may slightly differ from the one that includes lag selection
          due to different number of initial values used in test regression")
    print(summary(ur.df(data_aktien_zoo_select, type = adf_determ, lags = p-1, selectlags = "Fixed")))
    tseries::adf.test(data_aktien_zoo_select, alternative="stationary", k = p-1)
}

#   3) verwende R-Paket aTSA
    # enthält auch adf.test(), das p-Werte liefert
    # berechnet für alle drei Spezifikationen der deterministischen Terme
    #   - keine Konstante (drift), kein Trend
    #   - Konstante (drift), kein Trend (Fall A)
    #   - Konstante (drift), Trend (Fall B)
    # und allen Lagordnungen bis zu einer maximalen Lagordnung die ADF-Teststatistik.
    # Wird keine maximale Lagordnung angegeben, berechnet der Befehl automatisch
    # ein maximale Lagordnung.or all combinations of lag and trend specifications up to

if (!require(aTSA)) install.packages("aTSA")
library(aTSA)

aTSA::adf.test(as.vector(data_aktien_zoo_select)) #, nlag = adf_maxlags, output = TRUE)

# ------------------------------------------------------------------------
#                                Ende
# ------------------------------------------------------------------------
```

Listing A.14: ./R-code-ss26/ZOE_ss26_vorl_folien_R–14_SP500_ADF–Tests.R

B. R-Befehle für die Regressions- und Zeitreihenanalyse

B.1. Übersicht über verfügbare R-Befehle

Benötigte R-Pakete: `stats` (normalerweise geladen), `readxl`, `wooldridge`, `car`, `lmtest`, `moments`, `sandwich`, `zoo`, `dynlm` , `urca`, `tseries`.

Siehe hierzu auch Kleiber und Zeileis (2008), das eine sehr gute Einführung in R bietet.

Durchführen einer linearen Regression mittels `model_kq <- lm()` erstellt ein Regressionsobjekt, das Grundlage für die Befehle auf den folgenden Folien ist.

Anmerkung: Das Paket `dynlm` enthält den Regressionsbefehl `dynlm()`, der weitere Optionen zur erleichterten Programmierung dynamischer Modelle enthält, z.B. schnelle Eingabe der gelaggten Variablen, Saisondummies, Trendvariablen, etc... .

R-Befehl	Funktionsbeschreibung	R-Paket
	Schätz- und Prognoseoutput	
`print()`	einfaches gedrucktes Display	
`summary()`	Standard Regressionsoutput	
`coef()`	(oder `coefficients()`) extrahiert geschätzte Regressionsparameter	
`residuals()`	(oder `resid()`) extrahiert Residuen	
`fitted()`	(oder `fitted.values()`) extrahiert angepasste/gefittete Werte	
`anova()`	Vergleich von geschachtelten Modellen (nested models)	
`predict()`	Vorhersagen für neue Regressionswerte	
`confint()`	Konfidenzintervalle für Regressionskoeffizienten	
`confidenceEllipse()`	Konfidenzintervalle für Regressionskoeffizienten	`car`
`deviance()`	Residuenquadratsumme (RSS)	
`vcov()`	(geschätzte) Varianz-Kovarianzmatrix der Parameterschätzer	
`logLik()`	Log-Likelihood (unter der Annahme normalverteilter Fehler)	

Testen		
`hccm()`	Heteroskedastie-korrigierte Varianz-Kovarianzmatrix der Parameterschätzer; mit `type="hc0"` White-Varianz-Kovarianzmatrix, siehe **Methoden der Ökonometrie**	`car`
`coeftest()`	Standard-Regressionsoutput, ggf. mit heteroskedastie-robusten Standardfehlern	`lmtest`
`linearHypothesis()`	F-Test `test=c("F")` oder (asymptotischer) χ^2-Test `test=c("Chisq")`; mit `white.adjust=c(FALSE, TRUE, "hc0")` White-heteroskedastierobuste-Varianz-kovarianzmatrix	`car`
`lrtest()`	Likelihoodratio-Test, siehe **Weiterführende Fragen der Ökonometrie** oder **Fortgeschrittene Ökonometrie**	`lmtest`
`waldtest()`	Wald-Test, siehe **Weiterführende Fragen der Ökonometrie** oder **Fortgeschrittene Ökonometrie**	`lmtest`

Modellspezifikation		
`AIC()`	Informationskriterien einschließlich AIC, BIC/SC (unter der Annahme normalverteilter Fehler) - Beachte: Im Gegensatz zu EViews wird die geschätzte Parametervarianz als Parameter mitgezählt und nicht durch die Zahl der Beobachtungen dividiert, siehe **Methoden der Ökonometrie**	
`SelectCritEViews()`	Informationskriterien à la EViews, siehe **Methoden der Ökonometrie**	eigenes R-Programm, siehe Abschnitt B.2
`encomptest()`	Encompassing-Test zum Testen nicht geschachtelter Regressionsmodelle, siehe **Methoden der Ökonometrie**	`lmtest`
`jtest()`	J-Test zum Testen nicht geschachtelter Regressionsmodelle, siehe **Methoden der Ökonometrie**	`lmtest`
`acf()`	(graphische) Ausgabe der geschätzten Autokorrelationsfunktion einer Zeitreihe, verwende `acf(resid(model))` zur Betrachtung der Autokorrelationsfunktion in den Fehlern, siehe Abschnitt 4.1.1	
`pacf()`	analog zu `acf()`, jedoch Ausgabe der geschätzten partiellen Autokorrelationsfunktion	
`uc.df()`	**A**ugmented **D**ickey-**F**uller-Einheitswurzeltest zum Testen auf Stationarität einer Zeitreihe, siehe Abschnitt 9.2	`urca`

Modelldiagnose		
`plot()`	Graphiken zur Modellüberprüfung	
`resettest()`	RESET-Test zum Testen der funktionalen Form, siehe **Methoden der Ökonometrie**	`lmtest`
`jarque.test()`	Lomnicki-Jarque-Bera-Test zum Überprüfen normalverteilter Fehler, siehe **Methoden der Ökonometrie**	`moments`
`bptest()`	Breusch-Pagan-Test zum Testen auf Vorliegen von heteroskedastischen Fehlern, siehe **Methoden der Ökonometrie**	`lmtest`
`bgtest()`	Breusch-Godfrey-Test zum Testen auf Vorliegen von serieller Autokorrelation beliebiger Ordnung in den Fehlern, siehe Abschnitt 8.2	`lmtest`
`dwtest()`	Durbin-Watson-Test zum Testen auf Vorliegen von serieller Autokorrelation erster Ordnung in den Fehlern, siehe Abschnitt 8.1.4	`lmtest`
`whitetest()`	White-Test zum Testen auf Vorliegen von heteroskedastischen Fehlern, siehe **Methoden der Ökonometrie**	eigenes R-Programm, siehe **Methoden**

B.2. Eigene R-Funktionen

SelectCritEViews: R-Programm zur Berechnung der Modellselektionskriterien wie in EViews:

```
##############################################################################
#         Function SelectCritEViews
##############################################################################
# Function to compute model selection criteria like in EViews
# RT, 2011_01_26, 2017_05_18

SelectCritEViews <- function(model)
{
    n   <- length(model$residuals)
    k   <- length(model$coefficients)
    fitmeasure  <- -2*logLik(model)/n

    aic <- fitmeasure + k * 2/n
    hq  <- fitmeasure + k * 2*log(log(n))/n
    sc  <- fitmeasure + k * log(n)/n
    sellist <- list(aic=aic[1],hq=hq[1],sc=sc[1])
    return(t(sellist))
}

##############################################################################
#               End Function SelectCritEViews
```

```
################################################################################
```

Listing B.1: ../../R_code/SelectCritEViews.R

Mehr:

- Kurs **Programmieren in** R
- Kleiber und Zeileis (2008)
- Übersicht über verfügbare Pakete in R

  http://cran.r-project.org/web/views/

C. Versionsgeschichte

Diese Folien entsprechen den Folien vom April 2025 bis auf Änderungen auf folgenden Seiten: 1 - 4 (Organisation), 7 (Literaturangaben), 10, 11, 13, 14, 17, 82 (Datenaktualisierung), 104 (Hinzufügen Hinweis auf R-Programm), 120, 126, 151, 196, 197 (Ergänzungen und Korrekturen), 289, 301, 306 (Verweise auf R-Programme), ebenso neue Verweise in Appendix A: II, XVI, XVII.

Die Folien vom April 2025 entsprechen den Folien vom April 2022 bis

auf Änderungen auf folgenden Seiten: 6 - 7 (Literaturangaben), 9 - 10 , 12 - 13., 16 (jeweils Datenaktualisierung), 30, 34, 67-68 (Datenaktualisierung), 70, 77, 82, 108, 119, 120, 126, 128, 129, 166 (Bedingte Homoskedastie), 171 - 173, 174 (schwache Stationarität), 183 (Beweisinfo), 195, 196 (y_0), 200, 205 (jeweils zusätzliche Annahme), 220, 232 (bed. Erwartungswert), 290, 306.

Die Folien vom April 2022 entsprechen den Folien vom April 2014 bis auf folgende Änderungen:

- Alle empirischen Beispiele und Monte-Carlo-Simulationen sind in R durchgeführt. Die entsprechenden R-Programme befinden sich in Appendix A. Die Übersicht über R-Befehle befindet sich jetzt in Appendix B.

- Hinweise zu EViews wurden durch Hinweise zu R-Befehlen ersetzt.

- Die Daten zu den empirischen Beispielen wurden für S&P 500 und die deutschen realen BIP-Daten aktualisiert.

- Fehler wurden korrigiert.

Die Folien von 2014 entsprechen den Folien vom April 2010 mit folgenden Ausnahmen: i)Zu den Anwendungen werden im Appendix A Anwendungen jetzt neben EViews-Workfiles auch die entsprechenden Daten und Programme für R angeboten; ii)Fehlerkorrekturen; iii) teilweise Datenaktualisierungen; weiteren Fehlerkorrekturen seit April 2013.

Literaturverzeichnis

Brooks, C. (2008), *Introductory econometrics for finance*, 2nd edn., Cambridge University Press, Cambridge.

Brooks, C. (2019), *Introductory econometrics for finance*, 4th edn., Cambridge University Press, Cambridge.

Davidson, R., und J. MacKinnon (1993), *Estimation and Inference in Econometrics.*, Oxford University Press, URL http://www.oup.com/uk/catalogue/?ci=9780195060119. 285, 307

Deistler, M., und W. Scherrer (2018), *Modelle der Zeitreihenanalyse*, Mathematik Kompakt, Birkhäuser, doi:https://doi.org/10.1007/978-3-319-68664-6.

Diebold, F. (2007), *Elements of forecasting*, 4th edn., Thomson/South-Western.

Enders, W. (2015), *Applied Econometric Time Series*, 4th edn., Wiley.

Franses, P.-H. (1991), "Primary Demand for Beer in the Netherlands: An Application of the ARMAX Model Specification," *Journal of Marketing Research*, 28, 240–245. 26, 43, 46, 54

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press.

Kirchgässner, G., und J. Wolters (2008), *Introduction To Modern Time Series Analysis*, Springer, Berlin, [u.a.]. 173

Kirchgässner, G., J. Wolters, und U. Hassler (2013), *Introduction To Modern Time Series Analysis*, 2nd edn., Springer, Berlin, [u.a.].

Kleiber, C., und A. Zeileis (2008), *Applied Econometrics with R*, Springer, doi:10.1007/978-0-387-77318-6. LXX, LXXVI

Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer. 158

Lütkepohl, H., und M. Krätzig (eds.) (2004), *Applied Time Series Econometrics*, Cambridge University Press, Cambridge.

Lütkepohl, H. und Krätzig, M. (ed.) (2008), *Applied Time Series Econometrics*, Cambridge University Press.

Neusser, K. (2009), *Zeitreihenanalyse in den Wirtschaftswissenschaften*, 2nd edn., Teubner, Wiesbaden. 103

Neusser, K., und M. Wagner (2022), *Zeitreihenanalyse in den Wirt-*

schaftswissenschaften, 4th edn., Springer Gabler, Berlin Heidelberg. 173, 175

Newey, W., und K. West (1994), "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631–653, doi:doi:10.2307/2297912A. 276

Phillips, A. W. (1958), "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957," *Economica*, 283–299. 9

Schlittgen, Rainer und Sattarhoff, C. (2020), *Angewandte Zeitreihenanalyse mit R*, 4th edn., De Gruyter.

Wooldridge, J. M. (2009), *Introductory Econometrics. A Modern Approach*, 4th edn., Thomson South-Western, Mason. 5, 40, 47, 63, 65, 72, 78, 80, 102, 105, 110, 116, 148, 158, 164, 168, 169, 176, 179,